

### Verbesserung des Zugangs der Wissenschaft zu Mikrodaten: Pilotprojekt zum Aufbau eines Servicezentrums für Mikrodaten der GESIS bei ZUMA

Bohr, Jeanette; Janßen, Andrea; Lengerer, Andrea; Lüttinger, Paul; Schroedter, Julia H.; Wolf, Christof

Veröffentlichungsversion / Published Version  
Abschlussbericht / final report

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:  
GESIS - Leibniz-Institut für Sozialwissenschaften

#### Empfohlene Zitierung / Suggested Citation:

Bohr, J., Janßen, A., Lengerer, A., Lüttinger, P., Schroedter, J. H., & Wolf, C. (2007). *Verbesserung des Zugangs der Wissenschaft zu Mikrodaten: Pilotprojekt zum Aufbau eines Servicezentrums für Mikrodaten der GESIS bei ZUMA*. (ZUMA-Methodenbericht, 2007/05). Mannheim: GESIS-ZUMA. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-251033>

#### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

*ZUMA-Methodenbericht Nr. 2007/05*

**Verbesserung des Zugangs der  
Wissenschaft zu Mikrodaten. Pilotprojekt  
zum Aufbau eines Servicezentrums für  
Mikrodaten der GESIS bei ZUMA**

Jeanette Bohr, Andrea Janßen, Andrea Lengerer,  
Paul Lüttinger, Julia Schroedter, Christof Wolf

Juni 2007  
ISSN 1610-9953

GESIS-ZUMA  
Quadrat B2,1  
Postfach 12 21 55  
68072 Mannheim  
Telefon: 0621-1246-251  
Telefax: 0621-1246-100  
E-Mail: [christof.wolf@gesis.org](mailto:christof.wolf@gesis.org)

**GESIS**

Gesellschaft Sozialwissenschaftlicher  
Infrastruktureinrichtungen

**ZUMA**

Zentrum für Umfragen,  
Methoden und Analysen

**Forschungsverbund Datenservicezentren  
Verbesserung des Zugangs der Wissenschaft zu Mikrodaten  
Teilprojekt: Pilotprojekt zum Aufbau eines Servicezentrums  
für Mikrodaten der GESIS beim ZUMA**

BMBF Projekt  
Förderkennzeichen: 07SWF05B  
Gesamtbewilligungszeitraum 01.07.2003 – 31.12.2006

**Abschlussbericht**

GESIS – ZUMA  
PD Dr. Christof Wolf  
Postfach 12 21 55  
68072 Mannheim  
Tel.: +49 621 1246 153  
Telefax: +49 621 1246 100  
Email: [christof.wolf@gesis.org](mailto:christof.wolf@gesis.org)

## Zusammenfassung

Das Pilotprojekt zum Aufbau eines Servicezentrums für Mikrodaten der GESIS beim ZUMA hat zwischen dem 1. Juli 2003 und dem 31. Dezember 2006 drei Teilprojekte erfolgreich abgeschlossen. Ausgehend von den umfangreichen Erfahrungen der Abteilung Mikrodaten beim ZUMA wurden im Teilprojekt „Stärkung von Service, Beratung und Forschungsleistungen“ die Erschließung, Aufbereitung und Dokumentation aktueller Mikrozensus Scientific Use Files grundlegend neu konzipiert. Im Ergebnis werden diese Arbeiten heute effizienter durchgeführt und die resultierenden Produkte sind systematischer und umfangreicher. Heute finden Nutzer des Mikrozensus für alle zur Verfügung stehenden Scientific Use Files alle zur Arbeit notwendigen Informationen in strukturierter Form im Web-Angebot des GML. Außerdem wurden mehrere Workshops entwickelt, in denen das notwendige methodische und inhaltliche Wissen zum Umgang mit dem Mikrozensus vermittelt wird. Diesen Workshops stehen die in regelmäßigen Abständen durchgeführten Nutzerkonferenzen zur Seite, die dem Austausch zwischen Wissenschaft und amtlicher Statistik sowie der Präsentation neuester Forschungsergebnisse dienen.

Im Teilprojekt „Rückerschließung von Mikrozensus vor 1989“ mussten zunächst die rechtlichen Voraussetzungen für die Erstellung von Scientific Use Files für ältere Mikrozensus geprüft sowie eine begründete Auswahl von rückzuerschließenden Datensätzen erarbeitet werden. Im Anschluss daran wurde gemeinsam mit den Statistischen Ämtern des Bundes und der Länder ein Anonymisierungskonzept für die Scientific Use Files erarbeitet. Schließlich mussten für die erstellten Files die entsprechenden Dokumentationen und Tools erarbeitet und über das WWW bereitgestellt werden.

Im Teilprojekt „Aufbau eines Mikrodaten-Informationssystems“ wurde der Prototyp eines Informationssystems erarbeitet, in welchem prinzipiell alle für die Forschung relevanten Metadaten der einzelnen Erhebungen des Mikrozensus aufgenommen werden können. Das Informationssystem MISSY wurde anhand der Daten für die Erhebungszeitpunkte 1995 und 1997 umgesetzt. Die Leistungen von MISSY bestehen darin, dass alle zur Verfügung stehenden Informationen gebündelt und nach forschungsrelevanten Gesichtspunkten verknüpft sind. Durch die Bereitstellung einer Variablen-Zeitpunkt-Matrix und der Dokumentation von Variablenbrüchen unterstützt MISSY die Forschung bei zeitvergleichenden Analysen.

## Inhaltsverzeichnis

I	Kurzdarstellung	6
1	Stärkung von Service, Beratung und Forschungsleistungen	6
2	Rückerschließung von Mikrozensen vor 1989	7
3	Mikrodaten-Informationssystem MISSY	9
4	Mitarbeiterinnen und Mitarbeiter nach Teilprojekten	11
II	Eingehende Darstellung	12
1	Stärkung von Service, Beratungs- und Forschungsleistungen	12
1.1	Datenaufbereitung und Dokumentation	12
1.2	Service	14
1.2.1	Das WWW-Angebot des GML	14
1.2.2	Betreuung und Beratung von (Gast-)Wissenschaftlern	18
1.2.3	Nutzerberatungen	19
1.2.4	Wissensvermittlung	21
1.2.5	Nutzerkonferenz	22
1.3	Forschung mit Mikrozensus Scientific Use Files	23
1.3.1	Methodenforschung	23
1.3.2	Inhaltliche Forschung	25
2	Rückerschließung und Aufbereitung des Mikrozensus vor 1989	28
2.1	Rechtsproblematik	28
2.2	Datenauswahl	30
2.3	Erstellung und Prüfung der Scientific Use Files	31
2.4	Wissenschaftlicher Service	32
2.5	Daten der Volks- und Berufszählung 1970	33
2.6	Nutzen und Verwertbarkeit	35
2.7	Fortschritt bei anderen Stellen	35
2.8	Erfolgte/geplante Veröffentlichung des Ergebnisses	36

3	Mikrodaten-Informationssystem MISSY	38
3.1	Variablenbezogene Informationen	40
3.2	Zugänge zu variablenbezogenen Informationen	42
3.3	Studienbezogene Informationen	47
3.4	Usability-Studie	49
3.5	Technische Umsetzung	50
3.6	Nutzen und Verwertbarkeit	53
3.7	Fortschritt bei anderen Stellen	54
3.8	Erfolgte/geplante Veröffentlichung des Ergebnisses	54
III	Anlage: Veröffentlichungen, Vorträge und Veranstaltungen	56
IV	Anlage: Erfolgskontrollbericht zum Pilotprojekt zum Aufbau eines Servicezentrums für Mikrodaten	65
V	Kurzfassung: Berichtsblatt	67

## **I Kurzdarstellung**

Dieser Bericht informiert über die durchgeführten Arbeiten im Rahmen des Pilotprojekts zum Aufbau eines Servicezentrums für Mikrodaten der GESIS beim ZUMA (Projektlaufzeit 01.07.2003 – 31.12.2006). Unter den zahlreichen im Rahmen des Projekts durchgeführten Arbeiten, die im Folgenden eingehend beschrieben werden, sind die Dokumentation aktueller Mikrozensus bis zur Erhebung 2004, die Durchführung der 4. Mikrozensus Nutzerkonferenz, die Aufbereitung und Dokumentation von drei älteren Mikrozensus als Scientific Use Files und die Weiterentwicklung des Mikrodaten Informationssystems besonders hervorzuheben. Der vorliegende Bericht informiert detailliert über die einzelnen Arbeiten und Ergebnisse, die im Rahmen der Projekte erzielt wurden.

Das Pilotprojekt zum Aufbau eines Servicezentrums hatte drei Schwerpunkte:

- 1 Stärkung von Service, Beratung und Forschungsleistungen,
- 2 Rückerschließung von Mikrozensus vor 1989,
- 3 Aufbau eines Mikrodaten-Informationssystems.

### **1 Stärkung von Service, Beratung und Forschungsleistungen**

*Aufgabenstellung:* Die zentralen Aufgaben in diesem Bereich bestehen in der Erschließung, Aufbereitung und Dokumentation von aktuellen Mikrozensus Scientific Use Files. Den Nutzern des Mikrozensus werden alle zur Arbeit mit den Daten notwendigen Informationen über das Web-Angebot des GML zur Verfügung gestellt. Zur Vermittlung von notwendigem Wissen im Umgang mit den Daten des Mikrozensus werden außerdem Workshops zur Einführung und Vertiefung angeboten. Des Weiteren werden in regelmäßigen Abständen Nutzerkonferenzen organisiert, die dem Austausch zwischen Wissenschaft und amtlicher Statistik sowie der Präsentation neuester Forschungsergebnisse dienen. Zur Sicherung der Qualität dieses Services führen die Mitarbeiter des GML in unterschiedlichen Themengebieten eigene Forschungsarbeiten mit den Daten des Mikrozensus durch.

*Voraussetzungen:* Daten des Mikrozensus wurden erstmals im Jahr 1995 in Form von Scientific Use Files an die Wissenschaft weitergegeben. Die Aufbereitung und Dokumentation dieser Daten erfolgte von Beginn an im GML, so dass für die Weiterführung und den Ausbau dieser Arbeiten auf umfangreiche Expertise zurückgegriffen werden konnte. In der Vergangenheit konnten die Daten des Mikrozensus häufig nur zeitverzögert aufbereitet werden. Im Zuge der Stärkung des Servicebereichs konnten diese Rückstände aufgeholt und vorhandene Lücken geschlossen werden. Damit konnte die steigende Nachfrage der Wissenschaft nach aktuellen Daten bedient werden.

*Planung und Ablauf:* Der vor Projektbeginn bestehende Rückstand in der Aufbereitung der Mikrozensus Scientific Use Files 1999 und 2001 konnte planmäßig bis Ende 2004 aufgeholt werden. Die Mikrozensus Scientific Use Files 2002 und 2003 wurde im Laufe des Jahres 2005 komplett aufbereitet und dokumentiert, unmittelbar nachdem die Rohdaten vom Statistischen Bundesamt geliefert wurden. Die Arbeiten am Mikrozensus Scientific Use File 2004 wurde im Jahr 2006 abgeschlossen, in dem auch mit den Vorbereitungen für die Aufbereitung der Daten des Jahres 2005 begonnen wurde.

Ebenfalls wie geplant wurden während der Laufzeit des Projekts insgesamt vier Workshops zur Einführung in das Arbeiten mit dem Mikrozensus (ab Herbst 2004 nach einem neuen Konzept), ein thematischer Workshop und drei Methodenworkshops durchgeführt. Im Jahr 2005 wurde gemeinsam mit dem Statistischen Bundesamt eine Mikrozensus-Nutzerkonferenz organisiert. Als weiterer Meilenstein ist die Umsetzung eines neuen Konzepts für den Web-Auftritt des GML im Jahr 2005 zu nennen.

*Wissenschaftlicher und technischer Standard, an den angeknüpft wird:* Zur Aufbereitung und Dokumentation der Mikrozensus Scientific Use Files wurde an die vom GML erarbeiteten Standards angeknüpft. Mit den Arbeiten am Mikrozensus 2004 wurde erstmals damit begonnen, den Prozess der Datenaufbereitung und -dokumentation vom bisher konventionellen Verfahren schrittweise auf die Standards der Data Documentation Initiative (DDI) umzustellen.

*Zusammenarbeit mit anderen Stellen:* Die Prüfung, Aufbereitung und Dokumentation der Mikrozensus Scientific Use Files erfolgte in enger Absprache mit der Fachabteilung Mikrozensus des Statistischen Bundesamtes (Holger Breiholz). In einzelnen Fragen der Erstellung der Scientific Use Files erfolgte die Zusammenarbeit mit dem Forschungsdatenzentrum des Statistischen Bundesamtes (Heike Habla). Für die Workshops zur Einführung in das Arbeiten mit dem Mikrozensus wurden Referenten aus der Fachabteilung Mikrozensus sowie der Abteilung Mathematisch-statistische Methoden des Statistischen Bundesamtes gewonnen. Außerdem wurde die Mikrozensus-Nutzerkonferenz gemeinsam mit der Fachabteilung des Statistischen Bundesamtes organisiert. In Bezug auf die inhaltliche Forschung der Projektmitarbeiter fand eine Zusammenarbeit mit Frank Kalter (Universität Leipzig) sowie Thomas Klein (Universität Heidelberg) statt.

## 2 Rückerschließung von Mikrozensus vor 1989

*Aufgabenstellung:* Bis zum Projektbeginn lagen noch keine Scientific Use Files für Mikrozensususerhebungen vor 1989 vor. Die Aufgabe des Projektes „Rückerschließung“ im Service-



zentrum für Mikrodaten bestand in einem ersten Schritt darin, die Voraussetzungen für die Erstellung von Scientific Use Files zu schaffen, d.h. insbesondere die Möglichkeit zu prüfen, ob der §16 (6) des Bundesstatistikgesetzes von 1987 auch auf die Daten des Mikrozensus vor 1989 angewendet werden kann. Nach der Lösung der rechtlichen Problematik musste eine Auswahl der Daten stattfinden, für die Scientific Use Files erstellt werden sollten. In einem dritten Schritt wurde gemeinsam mit den zuständigen Stellen in den Statistischen Ämtern ein Anonymisierungskonzept für die Scientific Use Files erarbeitet. Viertens mussten dann für die erstellten Files die entsprechenden Dokumentationen erarbeitet und über das WWW bereitgestellt werden.

*Voraussetzungen:* Die Voraussetzungen zur Lösung der Problematik der Weitergabe von Mikrozensus Daten vor 1989 in der Form von Scientific Use Files waren zu Beginn des Projektes eher negativ einzuschätzen, da eine rückwirkende Anwendung des §16 (6) BStatG 1987 vom Statistischen Bundesamt bis zum Projektbeginn abgelehnt wurde.

*Planung und Ablauf:* In der ersten Projektphase wurde eine Rechtsposition erarbeitet, um das Problem der rückwirkenden Anwendung des BstatG1987 zu lösen. Danach erfolgte die Auswahl der Daten und die Konzeption von Anonymisierungsmaßnahmen für die ausgewählten Daten. Nach diesem Schritt erfolgte die Erstellung der Scientific Use Files sowie der entsprechenden Dokumentationen.

*Wissenschaftlicher und technischer Standard, an den angeknüpft wird:* Seit 1995 wird der Mikrozensus in Form von Scientific Use Files ab der Erhebung von 1989 an die Wissenschaft weitergegeben. Die hierfür erarbeiteten Anonymisierungsregeln waren auch Grundlage für die Erstellung der Scientific Use Files für die älteren Daten. Ebenso wurden die Standards für die Erstellung von Dokumentationen (Setups) etc., die für die neueren Daten existieren, auch für die älteren Daten zugrunde gelegt.

*Zusammenarbeit mit anderen Stellen:* Die Lösung der Rechtsproblematik erfolgte im Austausch mit dem Forschungsdatenzentrum des Statistischen Bundesamtes in Wiesbaden (Markus Zwick). Die Erstellung eines Anonymisierungskonzeptes für die Scientific Use Files der ausgewählten Mikrozensen erfolgte in enger Kooperation mit dem Forschungsdatenzentrum der Statistischen Landesämter am Landesamt für Statistik in Düsseldorf (Sylvia Zühlke).

### 3 Mikrodaten-Informationssystem MISSY

*Aufgabenstellung:* Das Mikrodaten-Informationssystem MISSY hat zum Ziel, das wissenschaftliche Arbeiten mit Mikrodaten durch inhaltlich orientierte Aufbereitung, Bereitstellung und Verknüpfung von Metadaten („Daten über Daten“) zu erleichtern. MISSY wurde anhand der Daten des Mikrozensus konzipiert und zunächst für die Erhebungszeitpunkte 1995 und 1997 umgesetzt. Die Leistungen von MISSY bestehen darin, dass erstens alle zur Verfügung stehenden Informationen gebündelt werden, so dass z.B. variablenspezifische Informationen auf einer Seite zusammengefasst sind. Zur Bündelung von Informationen tragen auch die nach sozial- und wirtschaftswissenschaftlichen Gesichtspunkten relevanten Verknüpfungen bei, die eine zweite Leistung von MISSY darstellen. Ziel dabei ist es einerseits, Zeit- und Arbeitsaufwand für den Forscher zu verringern, da er die Informationen nicht selbst aus verschiedenen Dokumenten zusammensuchen muss, und andererseits, den Forscher überhaupt auf die Existenz relevanter Informationen aufmerksam zu machen. Da es zum Mikrozensus als ursprünglich für die amtliche Statistik konzipierte Erhebung einige zusätzliche Dokumente wie z.B. einen Definitionskatalog gibt, besteht die Möglichkeit, dass solche Informationen nicht von den Forschern registriert werden, wenn sie nicht direkt von den inhaltlich relevanten Stellen auf Variablenebene verlinkt sind.

Eine dritte Leistung von MISSY betrifft den Zugang zu Informationen auf Ebene der Variablen: Ausgehend von unterschiedlichen Nutzerkenntnissen und –interessen werden verschiedene Möglichkeiten offeriert, um an die jeweils interessierende Detailinformation zu gelangen.

Viertens soll MISSY Analysen zur Vergleichbarkeit von Erhebungszeitpunkten erleichtern. Dazu wurde eine Matrix konzipiert, die alle Variablen für alle Erhebungszeitpunkte von 1989 – 2004 und zudem Informationen zur Vergleichbarkeit für die Jahrgänge 1995 und 1997 enthält.

*Voraussetzungen:* Die Basis für die inhaltliche Aufbereitung vorhandener Metadaten, für die Erstellung neuer Informationen und für die Verknüpfung von Informationen bilden jahrelange Erfahrungen des German Microdata Labs im Bereich Mikrodaten, insbesondere mit dem Mikrozensus. Im Rahmen seiner Beratungsfunktion für die Wissenschaft hat das German Microdata Lab einen guten Überblick über die Fragen und Probleme, auf die die Wissenschaftler im Umgang mit dem Mikrozensus stoßen, wie z.B. Probleme der intertemporalen Vergleichbarkeit. Auch durch die eigene Forschungstätigkeit seiner wissenschaftlichen Mitarbeiter verfügt das German Microdata Lab sowohl über die notwendigen Informationen als auch über sozialwissenschaftliches Know-how und Erfahrung im Umgang mit den

Daten des Mikrozensus. Da die technische Umsetzung innerhalb GESIS-ZUMA in enger Zusammenarbeit mit den wissenschaftlichen Mitarbeiterinnen realisiert wurde, ist MISSY auch in technischer Hinsicht direkt auf die Dokumentationsanforderungen des Mikrozensus zugeschnitten. Ausgangspunkt des Projekts bildet ein interner Projektbericht, der eine Bestandsaufnahme zu den Aufgaben und Herausforderungen des Informationssystems beinhaltet (Granato 2001).

*Planung und Ablauf:* Nach der Erstellung des oben erwähnten Projektberichts wurde ein Prototyp erstellt, an dem die technischen Leistungen des Systems erarbeitet wurden. Zeitgleich wurden einerseits die benötigten Metadaten aufbereitet bzw. erstellt und andererseits die Verknüpfungen und Zugänge erarbeitet und deren Praktikabilität überprüft. Nach diesem Schritt erfolgte die technische Umsetzung und die Festlegung des Layouts. Im Anschluss an die Fertigstellung der ersten MISSY-Version wurde eine Usability-Studie durchgeführt, deren Ergebnisse die Basis für eine weitere Optimierung des Informationssystems lieferte.

*Wissenschaftlicher und technischer Standard, an den angeknüpft wird:* Deutschsprachige Informationssysteme, die über einen ähnlichen Umfang und ein vergleichbares Funktionsreichtum wie MISSY verfügen, gibt es bislang – bis auf das für den Paneldatensatz SOEP angelegte SOEP-info – in den Wirtschafts- und Sozialwissenschaften kaum, so dass sich das Projekt fachlich am Informationssystem des British Household Panel Survey orientierte. Die inhaltliche Orientierung erfolgte zudem am bestehenden Bedarf an Metadaten und am sich in der täglichen Arbeit des GML herauskristallisierenden Dokumentationsbedarf.

*Zusammenarbeit mit anderen Stellen:* Die Erstellung von MISSY erfolgte im kontinuierlichen Austausch mit der Fachabteilung Mikrozensus des Statistischen Bundesamtes in Bonn, auf dessen Dokumente bei der Erstellung von Metadaten zurückgegriffen werden konnte. Eine Ausweitung der Kooperation mit anderen Stellen fand in der letzten Phase des Projektes im Rahmen der Usability-Studie statt. Dabei wurde die erste fertige Fassung von MISSY, die bereits online verfügbar war, verschiedenen Experten zur Evaluation vorgelegt: Dr. Maximilian Stempfhuber vom Informationszentrum Sozialwissenschaften in Bonn sowie Wendy L. Thomas vom Minnesota Population Centre haben in dieser Phase die Usability von MISSY begutachtet. Vorwiegend inhaltlich orientierte Anregungen kamen von den Mitarbeitern des Statistischen Bundesamtes (Gruppe VIII C - Mikrozensus, Arbeitskräftestichprobe, Haushalt und Familie) Holger Breiholz und Dr. Robert Herter-Eschweiler. Für das beantragte Folgeprojekt MISSY II, in dem die volle Implementierung des Informationssystems umgesetzt werden soll, wird GESIS-IZ, vormals Informationszentrum Sozialwissenschaft in Bonn, die technische Umsetzung übernehmen.

#### 4 Mitarbeiterinnen und Mitarbeiter nach Teilprojekten:

##### *Mikrozensus-Grundfiles (Laufzeit: 01.07.2003-31.12.2006)*

- Dr. Hossein Shahla, 100 % (01.07.2004-31.12.2006)
- Andrea Lengerer, 75 % (01.10.2004-30.06.2006)
- Julia H. Schroedter, 75 % (01.07.2004-30.06.2006)
- Yvonne Lechert, 50 %, (01.07.2006-31.12.2006)

##### *Mikrozensus Rückerschließung (Laufzeit: 01.07.2003-31.12.2006)*

- Dr. Paul Lüttinger, 75 % (01.07.2003-31.12.2006)
- Yvonne Lechert, 50 %, (15.02.2005-30.06.2006)

##### *Mikrodaten-Informationssystem: MISSY (Laufzeit: 01.07.2003-31.12.2006)*

- Dr. Andrea Janßen, 100 % (01.10.2004-31.12.2006)
- Jeanette Bohr, 100 % (15.10.2003-30.06.2006)

## II Eingehende Darstellung

Im Folgenden wird dargestellt, wie die Zuwendung verwendet wurde und welche Ergebnisse im Rahmen des Pilotprojekts zum Aufbau eines Servicezentrums der GESIS erzielt wurden.

### 1 Stärkung von Service, Beratungs- und Forschungsleistungen

Im Schwerpunkt Stärkung von Service, Beratungs- und Forschungsleistungen sind die folgenden Arbeiten erbracht worden:

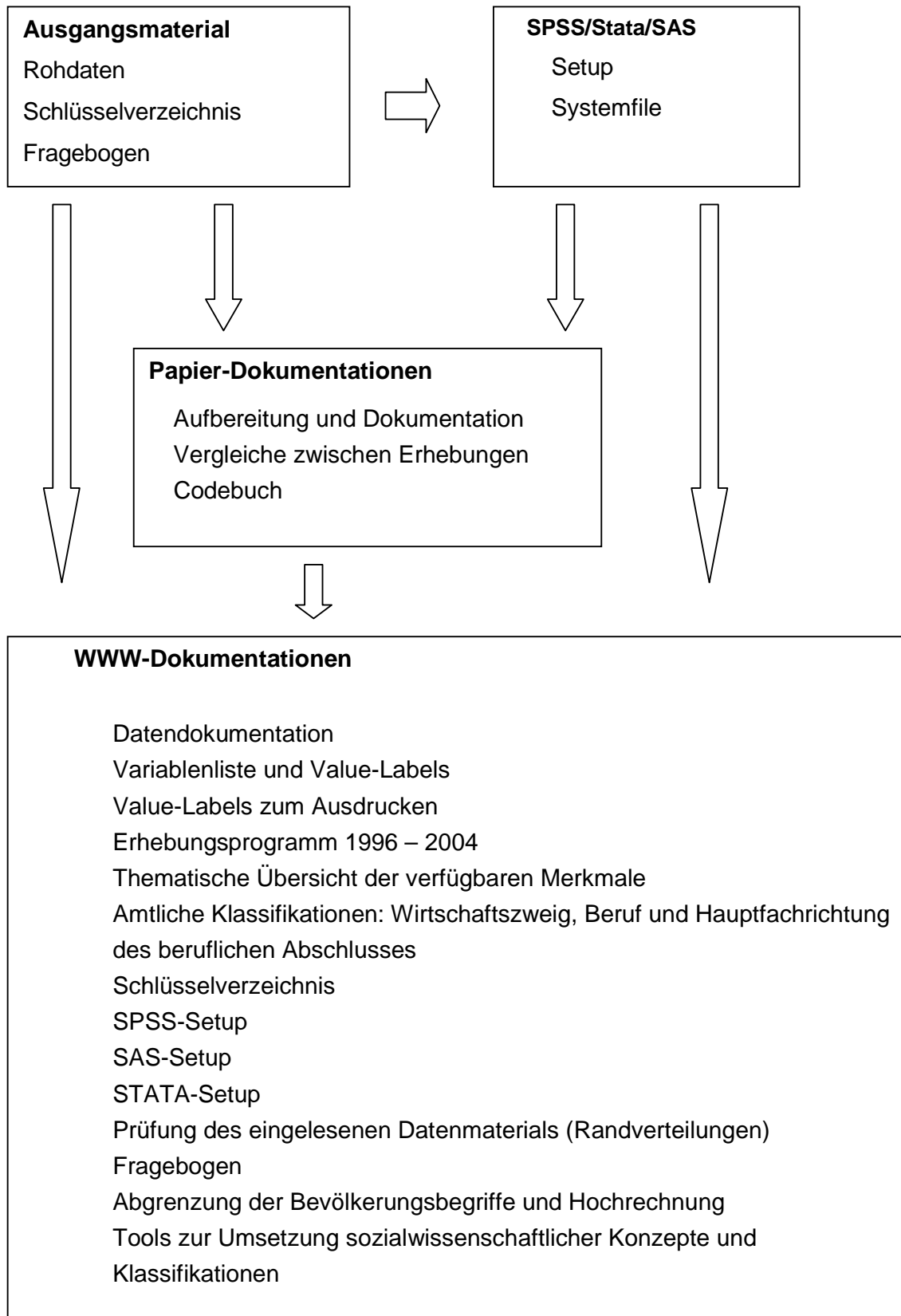
- **Datenaufbereitung und Dokumentation:** Die Mikrozensus 1999 bis 2004 wurden vollständig aufbereitet und dokumentiert. Die Arbeiten am Mikrozensus 2005 wurden im Berichtszeitraum aufgenommen.
- **Service:** Im Bereich Service wurden Nutzer der Mikrozensus Scientific Use Files beraten und unterstützt, Gastwissenschaftler betreut, Workshops zur Vermittlung von Kompetenzen bei der Nutzung der Scientific Use Files durchgeführt, die 4. Mikrozensus-Nutzerkonferenz organisiert und gemeinsam mit dem Statistischen Bundesamt durchgeführt, sowie das Webangebot des GML inhaltlich überarbeitet und ausgeweitet.
- **Forschung:** Bei der methodischen und inhaltlichen Forschung mit dem Mikrozensus lag der Fokus auf der Verbesserung des analytischen Zugangs, der Erleichterung und Erweiterung der Anwendungsmöglichkeiten der Scientific Use Files sowie auf eigener Forschung der Projektmitarbeiter mit den genannten Daten.

Die wichtigsten wissenschaftlich-technischen Ergebnisse der Aktivitäten im jeweiligen Arbeitsbereich werden im Folgenden genauer dargestellt.

#### 1.1 Datenaufbereitung und Dokumentation

Seit dem Bundesstatistikgesetz von 1987 können anonymisierte Individualdatenfiles der amtlichen Statistik durch die statistischen Ämter als so genannte faktisch anonymisierte Daten (§16, Abs. 6, BStatG) an die Forschung weitergegeben werden. Diese Daten liegen im „Rohdatenformat“ vor und sind – da sie nicht nach den gängigen Standards für sozialwissenschaftliche Daten aufbereitet, überprüft und dokumentiert sind – in dieser Form nicht ohne Weiteres für die wissenschaftliche Nutzung geeignet. Die Aktivitäten des Projektes im Bereich Datenaufbereitung und -dokumentation umfassen daher die Generierung von Einlese-routinen zur Aufbereitung des vom Statistischen Bundesamt gelieferten Rohdatenmaterials für verbreitete Statistiksoftware (SAS, SPSS, STATA), die Bereitstellung von Routinen zur Plausibilitätsprüfung sowie die Dokumentation des Aufbereitungsvorgangs (eine schematische Darstellung des Aufbereitungs- und Dokumentationsprozesses findet sich in der Übersicht 1).

## Übersicht 1: Prozess der Datenaufbereitung und WWW-Dokumentationen



Darüber hinaus werden in Methodenberichten die Spezifika des jeweiligen Mikrozensus sowie die Grenzen und Möglichkeiten der Vergleichbarkeit der Mikrozensusjahrgänge untereinander dargestellt, um die Nutzung der Daten für Längsschnittanalysen zu erleichtern.

In der Projektlaufzeit wurden die Mikrozensen 1999, 2001, 2002, 2003 und 2004 vollständig aufbereitet und dokumentiert. Mit den Arbeiten am Mikrozensus 2005 wurde im Berichtszeitraum begonnen. Die Arbeiten zur Aufbereitung und Dokumentation der Mikrozensen 1999 bis 2004 wurden im Februar 2006 abgeschlossen. Die Routinen zur Aufbereitung und wissenschaftlichen Dokumentation amtlicher Mikrodaten, die im Rahmen der regulären Serviceleistungen von GESIS-ZUMA entstanden sind, wurden in der Projektlaufzeit ergänzt, weiterentwickelt und qualitativ verbessert. Unter anderem wurden die bislang nur in SPSS erstellten Einleseroutinen und Mikrodaten-Tools zur Umsetzung sozialwissenschaftlicher Konzepte und Klassifikationen in SAS und STATA übersetzt, um Nutzern dieser Statistikprogramme – vor allem Wirtschaftswissenschaftlern – den Zugang zum Mikrozensus zu erleichtern. Bei der Dokumentation des Mikrozensus 2004 wurde erstmals damit begonnen, den Prozess der Datenaufbereitung und -dokumentation vom bisher konventionellen Verfahren schrittweise auf die Standards der Data Documentation Initiative (DDI) umzustellen. So wird das Mikrozensus Scientific Use File 2005 derzeit nach DDI-Standards aufbereitet und dokumentiert. Das Hauptziel dieser Umstellung ist es, zum einen den aktuellen internationalen Standards zur Aufbereitung und Dokumentation wissenschaftlicher Daten zu entsprechen und zum anderen die Anschlussfähigkeit an künftige Entwicklungen auf diesem Gebiet zu gewährleisten.

## 1.2 Service

### 1.2.1 Das WWW-Angebot des GML

Informationen, Dokumentationen und Tools zu den Mikrozensus Scientific Use Files werden von unseren Nutzern überwiegend auf der Homepage des GML gesucht. Um diesem Bedarf gerecht zu werden, wurde ein Konzept zur Verbesserung und Erweiterung des Internetauftritts erarbeitet. Das neue Konzept für den Web-Auftritt, das seit Mai 2005 umgesetzt ist, stellt die Übersichtlichkeit des Serviceangebotes in den Mittelpunkt und ermöglicht einen raschen Zugriff auf die einzelnen Dokumentationen und angebotenen Leistungen. Ein Großteil der von Nutzern an die Mitarbeiter des GML gerichteten Fragen lässt sich inzwischen mit einem Verweis auf die entsprechenden Stellen im Web zügiger und effizienter beantworten.

Die folgenden Tabellen geben einen Überblick über die Entwicklung der WWW-Seiten insgesamt sowie in einigen Teilbereichen für die Jahre 2003 bis 2006. Das WWW-Angebot

des GML ist in diesem Zeitraum kontinuierlich um mehrere hundert Seiten pro Jahr gewachsen und umfasst inzwischen annähernd 6.000 Seiten. Die Anzahl der so genannten Page-Views, d. h. die Anzahl der Aufrufe einzelner Seiten (die sich aus mehreren Dateien wie z. B. Frames und/oder Bildern zusammensetzen können), ist von 2003 bis 2005 kontinuierlich gestiegen. Der für 2006 sichtbare Rückgang resultiert aus einer technischen Änderung: So wurden bisher die Ausprägungen von Variablen (pro Erhebungsjahr etwa 300) auf jeweils einzelnen Seiten dargestellt, während dies seit der Bereitstellung des Mikrozensus 2004 im Jahr 2006 effizienter auf nur vier Seiten geschieht. Insgesamt erfolgten in diesem Jahr täglich rund 1.500 Zugriffe auf WWW-Seiten des GML (vgl. Tabelle 1).

**Tabelle 1: Anzahl der WWW-Seiten und -Zugriffe (GML-Angebot)<sup>a</sup>**

	Anzahl der Seiten	Page Views	Page Views pro Tag
2003	3.594	490.791	1.344
2004	4.415	716.106	1.961
2005 <sup>b</sup>	5.046	850.299	2.329
2006	5.759	550.549	1.508

<sup>a</sup> deutsche und englische Seiten

<sup>b</sup> Webadresse bis April 2005: /dauerbeobachtung/mikrodaten,  
ab Mai 2005: dauerbeobachtung/gml/

Der Großteil der Seiten beinhaltet Informationen über die Daten (ca. 90 % vgl. Tabelle 2), hierunter wiederum nimmt der Mikrozensus anteilig die zentrale Stellung ein. Über 80 % aller jährlichen Zugriffe auf das GML-Web erfolgen im Bereich Daten. Die überwiegende Mehrheit der Zugriffe gilt hierbei dem Mikrozensus (vgl. Tabelle 3).

**Tabelle 2: Anzahl der WWW-Seiten: /Daten**

	Anzahl der Seiten	% von Gesamt (Tab. 1)	Page Views	% von Ges. (Tab. 1)
2003	2.886	80,3	365.528	74,7
2004	3.606	81,7	541.393	75,6
2005	4.656	92,3	761.787	89,5
2006	5.170	89,7	475.349	86,3

**Tabelle 3: Anzahl der WWW-Seiten: /Mikrozensus**

	Anzahl der Seiten	% von Daten (Tab. 2)	Page Views	% von Daten (Tab. 2)
2003	2.098	72,7	272.075	74,7
2004	2.809	77,9	416.130	76,9
2005	3.707	79,6	595.503	78,2
2006	4.201	81,3	376.209	79,1

Ein wichtiges Serviceangebot stellen die so genannten Mikrodaten-Tools dar. Im Jahr 2006 wurde dieses Angebot um verschiedene Tools erweitert, was sich in einer Erhöhung der Seitenzahlen ausdrückt. Ebenso wurden bereits vorhandene Tools um bislang fehlende



Erhebungsjahre ergänzt. Die Seiten zu den Mikrodatentools wurden durchschnittlich etwa 50-mal pro Tag abgerufen.

**Tabelle 4: Anzahl der WWW-Seiten: /Mikrodaten-Tools**

	Anzahl der Seiten	Page Views	Page Views pro Tag
2003	106	14.537	39
2004	114	21.053	59
2005	127	19.206	53
2006	258	17.548	48

Tabelle 5 gibt die Seiten mit den meisten Zugriffen wieder. Insbesondere die Informationen zu den Daten, worunter auch ein Großteil der Methodenberichte fällt, aber auch Literaturhinweise sowie Mitteilungen unter Aktuelles befinden sich an der Spitze der Page Views.

**Tabelle 5: Hitliste der WWW-Seiten (nur HTML-Dokumente) 2006**

WWW-Seiten	Inhalte	Zugriffe
/daten/mz/index.htm	Informationen zum Mikrozensus	5.048
/daten/mz/mz_2004/index.htm	Informationen zum Mikrozensus 2004	2.047
/service/weitere_mikrodaten.htm	Beschreibung des Zugangs zu weiteren amtlichen Mikrodaten	1.763
/service/mz-suf-literatur/mz_literatur.htm	Liste der Veröffentlichungen auf Basis von Mikrozensus Scientific Use Files	1.753
/daten/mz/mz_2003/index.htm	Informationen zum Mikrozensus 2003	1.504
/f+e/publikationen/methodenberichte.htm	Methodenberichte (z.B. Dokumentation der Aufbereitung von Mikrozensus oder Beschreibung von Mikrodaten-Tools)	1.490
/daten/azbz/index.htm	Arbeitsstätten- und Berufszählungen	1.383
/daten/index.htm	Informationen zu den Daten insgesamt (z.B. Mikrozensus, Einkommens- und Verbrauchsstichprobe)	1.363
/mitarbeiter/index.htm	Informationen zu den Mitarbeitern	1.260
/daten/evs/index.htm	Informationen zur Einkommens- und Verbrauchsstichprobe	1.229
/daten/ddr/index.htm	Informationen zu Daten der amtlichen Statistik der DDR	1.207
/aktuelles/index.htm	Ankündigungen und Mitteilungen des GML	1.170
/daten/mz/mz-info/allg.htm	Allgemeine Informationen zu den Mikrozensus	1.100

**Tabelle 6: Downloads der Arbeits- und Methodenberichte 2006**

	Downloads
<b>Arbeitsberichte</b>	
Matthias Fleck, Georgios Papastefanou: Einkommens- und Verbrauchsstichprobe 1998 – Design und Methodik sowie Veränderungen gegenüber den Vorgängererhebungen. ZUMA-Arbeitsbericht Nr. 2006/01.	1.574
Bernhard Schimpl-Neimanns: Berufliche Ausbildungsverläufe bis zum Übergang ins Erwerbsleben: Analysen zur Stichprobenselektivität des Mikrozensuspanels 1996-1999. ZUMA-Arbeitsbericht Nr. 2006/02.	3.204
Bernhard Schimpl-Neimanns: Zur Datenqualität der Bildungsangaben im Mikrozensus. ZUMA-Arbeitsbericht Nr. 2006/03.	1.814
Bernhard Schimpl-Neimanns: Auszug aus dem Elternhaus: Ergebnisse des Mikrozensuspanels 1996-1999. ZUMA-Arbeitsbericht Nr. 2006/04.	1.198
<b>Methodenberichte</b>	
Lechert, Yvonne; Lüttinger, Paul: Mikrozensus Scientific Use File 1973. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/01.	183
Lechert, Yvonne; Lüttinger, Paul: Mikrozensus Scientific Use File 1976. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/02.	134
Lechert, Yvonne; Lüttinger, Paul: Mikrozensus Scientific Use File 1982. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/03.	186
Lengerer, Andrea; Shahla, Hossein: Mikrozensus Scientific Use File 2004 - Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/04.	459
Lengerer, Andrea; Boehle, Mara: Rekonstruktion von Bandsatzerweiterungen zu Haushalt, Familie und Lebensformen im Mikrozensus. ZUMA-Methodenbericht 2006/05.	213
Schroedter, Julia H.; Leim, Kathrin: Mikrozensus Scientific Use File 2004 - Vergleich mit dem vorhergehenden Erhebungszeitpunkt. ZUMA-Methodenbericht 2006/07.	620
Schroedter, Julia H.; Lechert, Yvonne; Lüttinger, Paul: Die Umsetzung der Bildungsskala ISCED-1997 für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensus 1976-2004. ZUMA-Methodenbericht 2006/08.	3.820
Lechert, Yvonne; Lüttinger, Paul; Schunck, Reinhard: Vergleich der Mikrozensus Scientific Use Files 1973 und 1976. ZUMA-Methodenbericht 2006/09.	360
Lechert, Yvonne; Lüttinger, Paul; Schunck, Reinhard: Vergleich der Mikrozensus Scientific Use Files 1976 und 1982. ZUMA-Methodenbericht 2006/10.	423
Lechert, Yvonne; Lüttinger, Paul: Volks- und Berufszählung (10 % Stichprobe). Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/11.	626
Lechert, Yvonne; Schroedter, Julia H.; Lüttinger, Paul: Die Umsetzung der Bildungsklassifikation CASMIN für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensus 1976-2004. ZUMA-Methodenbericht 2006/12.	175
Summe der Downloads 2006	14.989

Aus Tabelle 6 geht hervor, wie häufig die Arbeits- und Methodenberichte, die im Jahr 2006 erstellt wurden, von Nutzern heruntergeladen wurden. Mit 3.820 Downloads wurde der Methodenbericht zur Konstruktion der Bildungsskala ISCED mit Abstand am häufigsten nachgefragt, gefolgt vom Arbeitspapier zur Stichprobenselektivität des Mikrozensuspanels mit 3.204 Downloads. Insgesamt zeigen die Zahlen eine sehr aktive Nutzung der angebotenen

Berichte und Informationen durch die Scientific Community. Selbst Berichte zu speziellen Thematiken werden mit mehreren hundert Downloads relativ häufig abgerufen.

### 1.2.2 Betreuung und Beratung von (Gast-)Wissenschaftlern

Im Berichtszeitraum wurden insgesamt 24 Gastwissenschaftleraufenthalte im GML betreut (vgl. Tabelle 7). Die Aufenthalte werden zum einen von Wissenschaftlern außerhalb der Bundesrepublik Deutschland genutzt, an die aus rechtlichen Gründen keine Mikrozensus Scientific Use Files ausgehändigt werden dürfen. Zum anderen nehmen Wissenschaftler die Möglichkeit der Forschungsaufenthalte in Anspruch, um Daten auszuwerten, die sich im Datenbestand des GML befinden und nicht als Scientific Use Files zur Verfügung stehen (z.B. die Mikrozensus vor 1989, Volkszählung 1970, Mikrozensus-Zusatzerhebung 1971). Die Betreuung von Gastwissenschaftlern im GML umfasst die Bereitstellung von Daten und Variablenauszügen aus den Scientific Use Files, Hilfeleistungen bei der Nutzung der EDV-Infrastruktur sowie die inhaltliche Beratung bei der Datenauswertung .

**Tabelle 7: Übersicht über Aufenthalte von Gastwissenschaftlern im GML**

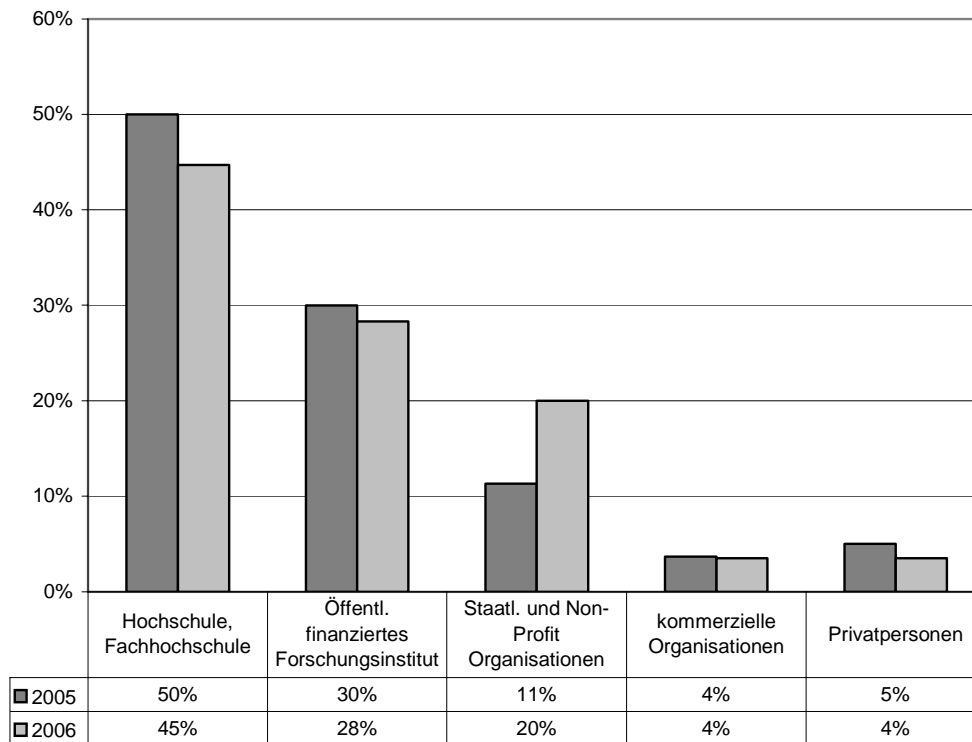
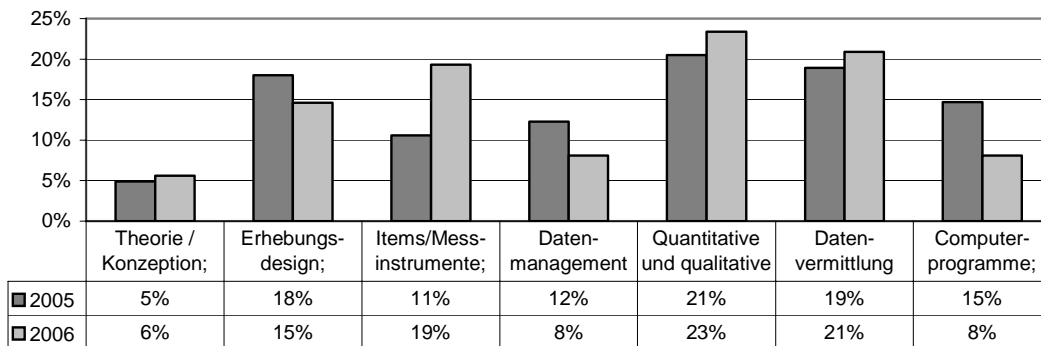
<b>Jahr, Name und Institution</b>	<b>Genutzte Daten</b>
<b>2006</b>	
Böhm, Sebastian (MPI für Demographische Forschung, Rostock) <sup>a</sup>	Volkszählung 1970
Kriwy, Peter (Universität Kiel)	Mikrozensus 2003
Radyakin, Sergiy (Technische Universität Darmstadt)	Mikrozensus 1963 - 1969
Yamai, Toshiaki (Ristumeikan Universität, Kyoto) <sup>a</sup>	Mikrozensus-Zusatzerhebung 1971
Waldinger, Fabian (London School of Economics & Political Science, London)	Mikrozensus 2002 - 2004
<b>2005</b>	
Böheim, René (Universität Linz)	Mikrozensus 1980 - 2002
Dümmmler, Kerstin (Technische Universität Dresden)	Mikrozensus 2001, 2002
Frietsch, Rainer (Fraunhofer Institut System- und Innovationsforschung, Karlsruhe)	Mikrozensus 1998 - 2002
Puhani, Patrick (Technische Universität Darmstadt) <sup>a</sup>	Mikrozensus 1963 - 1969
Radyakin, Sergiy (Technische Universität Darmstadt) <sup>a</sup>	Mikrozensus 1963 - 1969
Waldinger, Fabian (London School of Economics)	Mikrozensus 1989 - 2001
<b>2004</b>	
Dollmann, Jörg (MZES, Mannheim)	Mikrozensus 2002
Dümmmler, Kerstin (Technische Universität Dresden)	Mikrozensus 1989 - 2000

<b>Jahr, Name und Institution</b>	<b>Genutzte Daten</b>
Kesler, Christel (University of California Berkeley)	Mikrozensus 1996
Kohlberger, Jens (Universität Mannheim)	Mikrozensus 2000
Puhani, Patrick (Technische Universität Darmstadt) <sup>a</sup>	Mikrozensen 1963 - 1969
Trübswetter, Parvati (Technische Universität Darmstadt) <sup>a</sup>	Mikrozensen 1962 - 1969
Waldinger, Fabian (London School of Economics)	Mikrozensen 1989 - 2001
<b>2003</b>	
Frietsch, Rainer (Fraunhofer-Institut für Systemtechnik und Innovationsforschung, Karlsruhe)	Mikrozensen 1991, 1993
Kalter, Frank (Universität Mannheim)	Mikrozensus 1998
Kieffer, Annick (CNRS, LASMAS-IdL, Paris)	Mikrozensen 1973, 1985, 1993, 1998
Puhani, Patrick (Technische Universität Darmstadt) <sup>a</sup>	Mikrozensus 2000, Volkszählung 1970

a Im Rahmen des Teilprojekts „Rückerschließung“.

### 1.2.3 Nutzerberatungen

Weitere Aktivitäten des GML im Bereich Service betreffen die Betreuung und Unterstützung von Nutzern, die für ihre Forschungstätigkeiten die Daten der Mikrozensus Scientific Use Files über das Statistische Bundesamt bezogen haben und dabei auf Unterstützung in technischer und/oder inhaltlicher Hinsicht angewiesen sind. Dienstleistungen dieser Art erfolgen i.d.R. telefonisch oder per E-Mail. Die Anzahl der Anfragen hat sich gegenüber den Anfangsjahren deutlich reduziert, da mittlerweile über das WWW-Angebot des GML alle wichtigen Informationen zu den Daten, Verarbeitungshinweise etc. bereitgestellt werden. Im Jahr 2006 wurden von den Mitarbeitern des GML 82 (2005: 78) Beratungen mit einer Gesamtdauer von ca. 157 Stunden durchgeführt, wobei der überwiegende Teil der Anfragen (88 %) aus Deutschland kam sowie von Angehörigen von Hochschulen (45 %), öffentlich finanzierter Forschungseinrichtungen (28 %) oder sonstiger Non-Profit-Einrichtungen (20 %) (Abbildung 1). Die Verdoppelung der Beratungen bei staatlichen Organisationen von 2005 auf 2006 ist auf eine intensive Beratungstätigkeit z.B. bei Anonymisierungsprojekten der Ämter zurückzuführen. Ein Fünftel der Beratungen betraf Fragen zur Datenvermittlung, während sich der Rest der Anfragen auf mehrere Gebiete wie Datenanalyse oder Datenmanagement verteilte (Abbildung 2).

**Abbildung 1: Beratungen nach Art der Institution 2005, 2006****Abbildung 2: Beratung nach Themen 2005, 2006**

#### 1.2.4 Wissensvermittlung

Die Vermittlung von Kompetenzen zur Nutzung der Scientific Use Files in der empirischen Sozial- und Wirtschaftsforschung erfolgt im GML in Form von Einführungsworkshops. Diese Workshops haben über die Funktion des Wissenstransfers hinaus das Ziel, den Bekanntheitsgrad amtlicher Mikrodaten als Datenquelle für die sozialwissenschaftliche Forschung – insbesondere unter den Nachwuchswissenschaftlern – zu steigern und den Nutzerkreis dieser Daten zu erweitern.

In Kooperation mit dem Statistischen Bundesamt wurde erstmals im Frühjahr 2004 der Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Eine Einführung“ veranstaltet. Als Konsequenz aus der internen Evaluation des Workshops, der Teilnehmerbefragung und aus der gemeinsam mit dem Statistischen Bundesamt in 2003 durchgeführten Umfrage unter den Mikrozensusnutzern (Lüttlinger/Breiholz/Lechert 2003) wurde für die nachfolgenden Workshops ein neues Konzept erarbeitet und realisiert. So wurde die eintägige Veranstaltung mit Vorträgen zu Anlage, Inhalt und Durchführung des Mikrozensus um praktische Übungen mit den Daten ergänzt und auf zwei Tage ausgedehnt. Den Teilnehmerinnen und Teilnehmern wurde das Campusfile 1998 zur Verfügung gestellt, mit dem sie in den Übungen Datenaufbereitungs- und auswertungsarbeiten zu verschiedenen Themen durchführen konnten. So konnte nicht nur theoretisches, sondern auch anwendungsorientiertes Basiswissen und ein grundlegendes Verständnis für die Daten und deren Besonderheiten entwickelt werden. Die anschließende Evaluation des Workshops bestätigte die Vorteile des neuen Konzeptes (Schroedter/Lengerer 2005; Lengerer/Schroedter 2005). Im Rahmen dieser Veranstaltung wurde jedoch auch deutlich, dass das von den Forschungsdatenzentren der Statistischen Ämter erstellte Campusfile noch stärker an den konkreten Ausbildungsbedarf angepasst werden sollte. Dies betrifft insbesondere die Substichprobenziehung, den Merkmalsumfang, die Hochrechnungsfaktoren und ihre Dokumentation sowie die Dokumentation der Stichprobenziehung (Wirth/Schimpl-Neimanns 2004, in der Anlage).

Über diesen Einführungsworkshop hinaus wurden im Rahmen des ZUMA-Methodenworkshops im April und Oktober 2005 sowie im März 2006 drei Veranstaltungen zu speziellen Methoden angeboten, die für Analysen mit den Daten des Mikrozensus unerlässlich sind. Hierbei handelte es sich um die jeweils zweitägigen Workshops „Multivariate Analyse kategorialer Daten“. Sowohl der Einführungsworkshop als auch die Methodenworkshops werden überwiegend von Graduierten besucht, die mehrheitlich einen sozial- oder wirtschaftswissenschaftlichen Abschluss besitzen und an Hochschulen sowie an außeruniversitären Forschungsinstituten beschäftigt sind.

Im Herbst 2006 wurde weiterhin das Workshopangebot des GML um vertiefende, themenbezogene Workshops erweitert. Der erste Workshop dieser Art, der im Oktober 2006 zum Thema „Ausländische Bevölkerung im Mikrozensus“ angeboten wurde, wurde mit großem Interesse aufgenommen und war mit 20 Teilnehmern ausgebucht.

#### 1.2.5 Nutzerkonferenz

Im Oktober 2005 wurde die 4. Mikrozensus-Nutzerkonferenz mit dem Titel „Forschung mit dem Mikrozensus: Analysen zur Sozialstruktur und zum Arbeitsmarkt“ veranstaltet. Die seit 1998 von ZUMA gemeinsam mit dem Statistischen Bundesamt alle zwei Jahre durchgeführten Nutzerkonferenzen zum Mikrozensus wenden sich an Forscher, die mit Mikrozensus-Daten arbeiten oder daran interessiert sind. Ziele der Nutzerkonferenzen sind die Präsentation und Diskussion der auf Basis dieser Daten gewonnenen Forschungsergebnisse, wie auch die Förderung des Erfahrungsaustausches der Forscher untereinander und mit den statistischen Ämtern als Datenproduzenten.

Hierbei vereinen die Nutzerkonferenzen sowohl Service- als auch Forschungsaspekte. Der Faktor Service kommt insofern zum Tragen, als sich hier ein Forum bietet, bei welchem einerseits Forscherinnen und Forscher sich relativ einfach direkt mit den Experten des Statistischen Bundesamts wie auch des GML sowie anderen Datennutzern austauschen können. Zugleich profitiert das GML von der Diskussion mit den Datennutzern, da sich oftmals spezifische Probleme bei der Datennutzung herauskristallisieren, die als Indikatoren dafür herangezogen werden können, an welchen Stellen die Datennutzer mehr Hilfestellung benötigen. Gleichmaßen stellen die Nutzerkonferenzen wissenschaftliche Veranstaltungen dar, die sowohl Sozial- wie auch Wirtschaftswissenschaftlern aus wissenschaftlichen Instituten und Universitäten ein Forum für die Präsentation ihrer neuesten inhaltlichen oder methodischen Forschungsergebnisse mit amtlichen Mikrodaten bieten. Die inhaltlichen Schwerpunkte der Nutzerkonferenz im Oktober 2005 lagen auf bildungs- und genderspezifischen Aspekten des Arbeitsmarktes sowie den Potenzialen wie auch möglichen Restriktionen des Mikrozensus als Datenquelle für die Analyse spezifischer sozialer Lagen. Im Vordergrund standen zeitvergleichende Analysen; daneben wurden in den Beiträgen jedoch auch spezifische methodische Aspekte bei der Verwendung von Mikrozensusdaten in der Sozial- und Arbeitsmarktforschung diskutiert. Der methodische Schwerpunkt dieser Konferenz war auf die Mikrozensus-Paneldaten ausgerichtet, die inzwischen als Scientific Use File verfügbar sind. Hierbei präsentierten die Mitglieder des Projektes Methodenverbund „Aufbereitung und Bereitstellung des Mikrozensus als Panelstichprobe“ ihre bisherigen Forschungsergebnisse erstmals in umfassender Weise der Forschungsöffentlichkeit. Die auf den bisherigen Nutzer-

konferenzen gehaltenen Vorträge sind – soweit die Zustimmung der Autoren vorliegt – der breiten Forschungsöffentlichkeit über die WWW-Seiten des GML zugänglich.

### 1.3 Forschung mit Mikrozensus Scientific Use Files

Neben den Serviceleistungen und der Wissensvermittlung wurden in den Projekten methodische und inhaltliche Forschungsarbeiten durchgeführt. Diese Forschungsarbeiten erfolgten auf der Grundlage des Datenbestandes, zu dem der wissenschaftliche Service im GML angeboten wird und ist die notwendige Voraussetzung für eine kompetente Nutzerberatung.

#### 1.3.1 Methodenforschung

Im Bereich der methodischen Forschung wurden verschiedene Arbeiten durchgeführt. Es wurden drei neue Tools erstellt sowie vorhandene Tools fortgeschrieben.

*Umsetzung der „CASMIN“- und ISCED-Bildungsklassifikationen mit den Mikrozensus 1976-2004, der Volkszählung von 1970 und der Mikrozensus-Zusatzerhebung von 1971:* Um den Nutzern von Daten der amtlichen Statistik den Umgang mit den bildungsstatistischen Merkmalen zu erleichtern, wurden zwei Bildungsskalen umgesetzt, die sowohl in nationalen Untersuchungen als auch im internationalen Vergleich in den letzten Jahren zunehmend an Bedeutung gewonnen haben. Dabei handelt es sich um die von Sozialwissenschaftlern entwickelte CASMIN-Klassifikation sowie die von der UNESCO konstruierte ISCED-Klassifikation, die für die Mikrozensus 1976 bis 2004 (Scientific Use Files), die Volkszählung 1970 sowie die Mikrozensus-Zusatzerhebung 1971 umgesetzt wurden. Da sich die Kategorien der Bildungsangaben in diesen Erhebungen z.T. erheblich unterscheiden und eine intertemporale Vergleichbarkeit angestrebt wurde, wurden die Umsetzungen der Bildungsklassifikation ausführlich dargestellt. Eine Schwierigkeit bei der Konstruktion einer Klassifikation, in die der allgemein bildende und der berufliche Abschluss gemeinsam eingehen, stellen fehlende Angaben zu einer dieser beiden Variablen dar. Um dieses Problem zu lösen, wurden Regeln für den Umgang mit fehlenden Werten erarbeitet und diese auf ihre empirische Tauglichkeit hin überprüft.

Lechert, Yvonne, Julia H Schroedter und Paul Lüttinger, 2006: Die Umsetzung der Bildungsklassifikation CASMIN für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensus 1976-2004. ZUMA-Methodenbericht 2006/12

Schroedter, Julia H., Yvonne Lechert und Paul Lüttinger, 2006: Die Umsetzung der Bildungsskala ISCED-1997 für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensus 1976-2004. ZUMA-Methodenbericht 2006/08.



*Zum Konzept der Lebensformen und Familien im Mikrozensus:* Für die Bearbeitung familienwissenschaftlicher Fragestellungen sind im Mikrozensus vielfältige Angaben enthalten. Dabei handelt es sich vor allem um so genannte Bandsatzerweiterungen und Typisierungen, d.h. um generierte Variablen, für deren Auswertung die in der amtlichen Statistik verwendeten Definitionen von Haushalt, Familie und Lebensformen und die darauf aufbauenden Typologien von grundlegender Bedeutung sind. Um hier den Nutzern den Zugang zu erleichtern, sind diese in einem Arbeitsbericht vergleichend dargestellt und diskutiert. Der Schwerpunkt lag dabei auf einer Gegenüberstellung des traditionellen Familienkonzeptes und dem neuen Konzept der Lebensformen, das ab dem Mikrozensus 1996 umgesetzt ist.

Lengerer, Andrea, Jeanette und Andrea Janßen, 2005: Haushalte, Familien und Lebensformen im Mikrozensus – Konzepte und Typisierungen. ZUMA-Arbeitsbericht 2005/05.

Lengerer, Andrea, Andrea Janßen und Jeanette Bohr, 2007: Familiensoziologische Analysepotentiale des Mikrozensus. Erscheint in: Zeitschrift für Familienforschung, 19, Heft 2

*Rekonstruktion von Bandsatzerweiterungen zu Haushalt, Familie und Lebensformen im Mikrozensus:* Der Mikrozensus hält neben Informationen zu Bevölkerung und Arbeitsmarkt auch für familienwissenschaftliche Zwecke relevante Angaben bereit. Allerdings werden nur wenige haushalts- und familienbezogene Merkmale direkt erhoben: Lediglich acht der insgesamt rund 150 Fragen sind diesbezüglich verfasst. Das familiensoziologische und -demographische Analysepotential wird erst durch die so genannten Bandsatzerweiterungen ausgeschöpft. Aus der Kombination der im Mikrozensus direkt erhobenen Merkmale werden von den statistischen Ämtern zahlreiche Variablen im Nachhinein generiert, mit Hilfe derer schließlich umfangreiche familienbezogene Auswertungen möglich sind. Durch das seit 1996 eingeführte Konzept der Lebensformen hält der Mikrozensus neben Bandsatzerweiterungen zu Haushalt und Familie auch solche für Lebensgemeinschaften bereit. Damit sind nun differenzierte Analysen auf drei Ebenen möglich. In dem Methodenbericht 2006/05 werden die in den Bandsatzerweiterungen enthaltenen Informationen transparent und nachvollziehbar dargestellt. Dieses geschieht anhand der Rekonstruktion der bereits generierten Variablen zu Haushalt, Familie und Lebensformen mittels zeilenübergreifender Operationen in SPSS. Auf diese Weise wird auch aufgezeigt, wie für eigene Fragestellungen relevante Variablen selbst berechnet werden können, sofern sie nicht von den statistischen Ämtern zur Verfügung gestellt werden. Dieser Bericht wurde um ein Papier ergänzt, in dem gezeigt wird, wie diese Berechnungen mit Stata durchgeführt werden können ().

Lengerer, Andrea; Boehle, Mara, 2006: Rekonstruktion von Bandsatzerweiterungen zu Haushalt, Familie und Lebensformen im Mikrozensus. ZUMA-Methodenbericht 2006/05.

John, Kristina, 2006: Rekonstruktion von Bandsatzerweiterungen mittels zeilen- bzw. satz-übergreifender Operationen in STATA. ZUMA.

*Fortschreibung vorhandener Mikrodaten-Tools:* Neben der Entwicklung neuer Tools wurden im Berichtszeitraum die für ältere Mikrozensusjahrgänge erstellten Mikrodaten-Tools zur Umsetzung der Magnitude-Prestige-Skala, des internationalen sozioökonomischen Indexes ISEI und der Berufsklassifikation nach Blossfeld für die aktuellen Mikrozensus fortgeschrieben. Außerdem wurden teilweise vorhandene Lücken in der Bereitstellung von Tools für einzelne Erhebungsjahre geschlossen.

### 1.3.2 Inhaltliche Forschung

Die inhaltliche Forschung der Mitarbeiterinnen und Mitarbeiter im Bereich Stärkung Service, Forschung und Beratung richtet sich insbesondere auf die folgenden Themen.

*Determinanten der Teilzeittätigkeit von Frauen und Männern:* Ziel der Forschung in diesem Bereich ist es, das analytische Potenzial der Scientific Use Files des Mikrozensus für die Bearbeitung von inhaltlichen Fragestellungen bezogen auf das Erwerbsleben von Frauen und Männern zu nutzen und die wenig beachteten Aspekte dieser Thematik aufzuzeigen. Der Ausgangspunkt der Analysen ist die in allen empirischen Studien zu beobachtende Überrepräsentation der Frauen im Teilzeitsektor. Differenziertere Analysen zeigen jedoch in dieser Hinsicht grundlegende Unterschiede zwischen Frauen mit unterschiedlichen „Lebensformen“ und lassen die weit verbreitete These „Teilzeittätigkeit sei weiblich“ zumindest relativieren.

*Wandel von Lebensformen:* Der Wandel partnerschaftlicher Lebensformen wird sehr unterschiedlich gedeutet, ist empirisch aber nur unzureichend untersucht. Auf Basis national repräsentativer Daten mangelt es zum einen an Betrachtungen über lange Zeiträume und verschiedene Kohorten hinweg. Zum anderen ist über den bekannten Rückgang der Heiratsneigung und der damit einhergehenden Ausbreitung nichtehelicher Lebensgemeinschaften hinaus offen, ob sich auch die Neigung zum Eingehen fester partnerschaftlicher Beziehungen verändert hat.

Zu dieser Thematik wurde in 2006 ein Beitrag verfasst (Lengerer/Klein 2007). Darin wird der Wandel partnerschaftlicher Lebensformen über einen Zeitraum von mehr als vier Jahrzehnten hinweg untersucht. Als zentrales Kriterium zur Bestimmung der partnerschaftlichen Lebensform wird das (Nicht-) Vorhandensein eines Partners im Haushalt herangezogen. Differenziert wird zwischen Ehen, nichtehelichen Lebensgemeinschaften von Personen unterschiedlichen und gleichen Geschlechts sowie der Beziehungslosigkeit. Datengrundlage sind verschiedene

Mikrozensus-Erhebungen aus den Jahren 1962 bis 2004, die insgesamt Angaben aus über 12,6 Millionen Personen-Interviews enthalten. Nichteheliche Formen des Zusammenlebens werden im Mikrozensus zwar erst seit 1996 mit einer direkten Frage erfasst, für die Zeit davor sind aber valide Approximationen möglich. Entgegen weit verbreiteter Auffassungen kann ein Trend zur Abkehr von festen, verbindlichen partnerschaftlichen Beziehungen nicht festgestellt werden. Auch der bekannte Rückgang der Heiratsneigung und die Ausbreitung nichtehelicher Formen des Zusammenlebens können nicht pauschal im Sinne einer Pluralisierung gewertet werden. Vielmehr vollzieht sich der Wandel je nach Altersgruppe auf unterschiedlichem Niveau und teilweise auch in eine unterschiedliche Richtung. Ein Vergleich der Lebensverläufe verschiedener (Quasi-) Kohorten zeigt, dass partnerschaftliche Bindungen zunehmend später und bis zum mittleren Erwachsenenalter auch seltener eingegangen werden. Im höheren Alter nimmt hingegen der Anteil derer, die in Ehe oder Lebensgemeinschaft leben, zu, was vor allem auf die Veränderung der Alters- und Geschlechterstruktur der älteren Bevölkerung zurückzuführen ist.

*Migration und Integration:* Sowohl in der wissenschaftlichen als auch in der gesellschaftlichen Debatte um soziale Integration von Migranten gewinnt der sozialräumliche Bezug an Bedeutung. Die residentielle Segregation wird dabei oftmals als eine der Ursachen oder zumindest als verstärkendes Element für ein Scheitern der Integration von Migranten angesehen. Mit Daten des Mikrozensus wird untersucht, ob ein Zusammenhang zwischen der sozialen Integration und der residentiellen Segregation von Zugewanderten und deren Folgegeneration(en) zu beobachten ist. Die Möglichkeiten regionaler Analysen auf Basis des Mikrozensus sind aufgrund der Anonymisierungsanforderungen sehr eingeschränkt. Räumliche Aspekte können jedoch auf Grundlage der ausgewiesenen Auswahlbezirke aufgegriffen werden. Anhand dieser Information wird ethnische Segregation in Westdeutschland für die größten Migrantengruppen der ehemaligen Anwerbeländer im Zeitverlauf betrachtet. Erste Analysen zeigen, dass die räumliche Segregation Zugewanderter von 1976 bis 1982 zugenommen hat, während zwischen den Jahren 1996 und 2004 ein abnehmender Trend zu beobachten ist. Zudem zeigt sich, dass sich der Grad an ethnischer Segregation kaum nach Generationenzugehörigkeit unterscheidet, zumindest nicht, wenn man die mittleren Anteile an ausländischer Bevölkerung im Auswahlbezirk als Referenz betrachtet. Bei der ausschließlichen Betrachtung von ethnisch hoch segregierten Bezirken (Ausländeranteil mindestens 40 Prozent) lässt sich dagegen ein Unterschied zwischen erster und zweiter Migrantengeneration feststellen, wenn auch die Nationalitäten sehr viel größere Unterschiede aufweisen: So ist die Wahrscheinlichkeit bei Angehörigen der türkischen oder griechischen

Bevölkerung besonders hoch, in einem hoch segregierten Auswahlbezirk zu leben, während sie für Spanier und Italiener geringer ausfällt. Wird Bildungsniveau und Einkommensarmut berücksichtigt, zeigt sich ein deutliches Muster: Bei Einkommensarmut und niedriger Bildung bzw. Ausbildung wächst die Wahrscheinlichkeit, in einem Auswahlbezirk mit einem hohen Anteil an Ausländern zu leben. Dies gilt sowohl für Migranten als auch für Deutsche; allerdings bleiben auch unter Berücksichtigung dieser Faktoren deutliche Unterschiede zwischen den Nationalitäten bestehen (Andrea Janßen, Julia Schroedter).

*Ehen von Migranten in Westdeutschland:* Interethnische Ehen zählen seit langem zu den zentralen Indikatoren der Integration, da sie Auskunft über dauerhafte soziale Interaktionen von Gesellschaftsmitgliedern über Gruppengrenzen hinweg geben. Der Beitrag beschäftigt sich daher vornehmlich mit Ehen, die zwischen Migranten der ehemaligen Anwerbeländer und Deutschen bestehen. Neben einer deskriptiven Analyse werden erklärende Faktoren für die bestehenden Heiratsmuster – v.a. hinsichtlich des Bestehens binationaler Ehen – aufgezeigt. Den theoretischen Hintergrund bilden dabei Theorien der Partnerwahl sowie das Stufenmodell der Integration von Esser. Eine wichtige Annahme besteht darin, dass die strukturelle Integration einen entscheidenden Einfluss auf die soziale Integration ausübt. D.h. je stärker Zuwanderer in die Funktionssysteme der Aufnahmegesellschaft integriert sind, desto öfter führen sie eine Ehe mit einem Inländer bzw. einer Inländerin.

Da Zuwanderern neben dem deutschen Heiratsmarkt teilweise noch der Heiratsmarkt ihres Herkunftslandes zur Verfügung steht, wird zudem dargestellt, inwieweit die Option der transnationalen Ehe von den verschiedenen Nationalitäten genutzt wird und welche Einflussfaktoren dabei ggf. bedeutsam sind.

Die Analysen basieren auf den Mikrozensus der Erhebungsjahre 1976, 1982, 1991, 1995 und 1996-2004, sodass auch der Wandel über die Zeit in den Heiratsmustern abgebildet werden kann (Julia Schroedter).

## 2 Rückerschließung und Aufbereitung des Mikrozensus vor 1989

Der Gründungsausschuss des Rates für Sozial- und Wirtschaftsdaten hatte mit der Einrichtung des Servicezentrums für Mikrodaten bei ZUMA (Zentrum für Umfragen, Methoden und Analysen) auch die Empfehlung an das BMBF ausgesprochen, mit besonderer Dringlichkeit die Rückerschließung und Aufbereitung älterer Mikrozensus zu bearbeiten. Damit war der Zeitraum vor 1989 gemeint, für den noch keine Mikrozensusdaten in Form von Scientific Use Files für die Wissenschaft zur Verfügung standen. Damit wurde jedoch nicht nur einer „formalen“ Forderung nach sozialhistorischen Daten entsprochen, sondern auch dem konkreten Bedarf der Forschung nach älteren Daten. So äußerten bei einer Befragung, die ZUMA und das Statistische Bundesamt im Mai 2003 unter den Nutzern der Scientific Use Files durchführten, 40 Prozent der Befragten Interesse an weiteren Mikrozensusdaten für den Zeitraum vor 1989.

Die Aufgabe des eigens dafür eingerichteten Teilprojektes „Rückerschließung“ im Servicezentrum für Mikrodaten bestand folglich darin, auch für die Mikrozensuserhebungen vor 1989 für die Wissenschaft so genannte Scientific Use Files zu erschließen und einen entsprechenden wissenschaftlichen Service zu den Daten für die Profession anzubieten. Die Notwendigkeit eines eigenen Projektes lag primär darin begründet, dass bis zum Projektbeginn die Mikrozensusdaten von Erhebungen, die vor 1989 stattgefunden haben, nur in absoluter anonymisierter Form von den Ämtern an die Wissenschaft weitergegeben wurden. Dies beruhte im wesentlichen darauf, dass eine rückwirkende Anwendung des §16 (6) BStatG 1987 vom Statistischen Bundesamt generell nicht als Option in Erwägung gezogen wurde, sondern von der weiteren Gültigkeit des Bundesstatistikgesetzes von 1980 für die älteren Daten ausgegangen wurde.

### 2.1 Rechtsproblematik

Das Erreichen dieses Projektzieles war folglich eng mit der Frage verknüpft, ob das Bundesstatistikgesetz von 1987 (BStatG 87), das die Grundlage für die Weitergabe von Einzeldaten an die Wissenschaft in Form von Scientific Use Files bildet, auch auf den Zeitraum vor 1987 angewendet werden konnte. Die Erhebungen der amtlichen Statistik in der Bundesrepublik werden insbesondere durch das Bundesstatistikgesetz geregelt. Seit der Gründung der Bundesrepublik waren verschiedene Bundesstatistikgesetze gültig, das Statistik Gesetz von 1953 (StatGes 1953), das BStatG 1980 und das BStatG 1987. In diesen Gesetzen wird auch die Weitergabe von Daten an die Wissenschaft geregelt. Während das StatGes von 1953, das bis 1980 gültig war, keine wesentlichen Einschränkungen bei der Weitergabe von Daten an die

Wissenschaft vorsah, kam es mit der Neufassung des BStatG1980 zu einer wesentlichen Verschärfung der Weitergabebestimmungen die dazu führte, dass die Daten nur noch in absolut anonymer Form an die Wissenschaft übermittelt werden konnten. Bei der Weitergabe von Daten an die Wissenschaft, die unter der Gültigkeit des Statistikgesetzes von 1980 erhoben wurden, ergibt sich somit das Problem des sog. „Rückwirkungsverbotes“ von Gesetzen. Damit ist die (juristische) Frage gemeint, inwieweit die Bestimmungen des BStatG 1987 zur Weitergabe von Daten in Form von Scientific Use Files auch auf die Weitergabe von solchen Daten angewendet werden können, die unter der Geltung des BStatG80 erhoben wurden, da diese Daten einem strengeren Datenschutz unterlagen und der durch die Anwendung des BStatG87 sozusagen aufgehoben werden würde. Diese Frage war insofern von fundamentaler Bedeutung, als das Statistische Bundesamt bis zu diesem Zeitpunkt die Möglichkeit einer rückwirkenden Anwendung des BStatG 1987 unter Hinweis auf das Rückwirkungsverbot ausschloss und keine Möglichkeit sah, Daten, die vor 1989 erhoben wurden, als Scientific Use File der Wissenschaft zur Verfügung zu stellen.

In einem Arbeitspapier erarbeitete das German Microdata Lab zu dieser Frage eine eigene Rechtsposition (Lüttinger/Wirth 2004: Zur Weitergabe von älteren amtlichen Mikrodaten an die Wissenschaft. Allgemeines Statistisches Archiv 88). In diesem Papier wird ausführlich begründet, weshalb eine rückwirkende Anwendung des BStatG 1987 möglich sein sollte. Insbesondere für den Zeitraum bis 1980 (Gültigkeit StatGes 1953) sollte dies nach Auffassung der Autoren rechtlich unbedenklich sein. Das StatGes 1953 sah keinen expliziten Ausschluss der Wissenschaft als Datenempfänger vor, so dass es zu sehr „wissenschaftsfreundlichen Regelungen“ hinsichtlich der Datenweitergabe an die Wissenschaft und damit verbundener Anonymisierungsregeln kam. Auch für den Zeitraum 1980 bis 1987 (Gültigkeit BStatG 1980) sollte nach Meinung der Autoren eine Rückwirkung des BStatG 1987 möglich sein, zum einen aufgrund des besonderen Status, den der Gesetzgeber der Wissenschaft zuerkannt hat und der auch im Volkszählungsurteil von 1983 explizit anerkannt wurde. Zum anderen aber aufgrund der Tatsache, dass der Mikrozensus eine Pflichterhebung darstellt und somit kein schutzwürdiges Vertrauen - eine Voraussetzung für das Rückwirkungsverbot - aufgebaut werden konnte.

Anfang Februar 2004 wurde diese Rechtsauffassung des German Microdata Lab mit dem Statistischen Bundesamt diskutiert. Nachdem das Statistische Bundesamt bis zu diesem Zeitpunkt die Möglichkeit einer rückwirkenden Anwendung des BStatG 1987 ausschloss, kam nun auch die Arbeitsgruppe für Rechtsfragen der Forschungsdatenzentren zu dem Schluss, dass das Bundesstatistikgesetz von 1987 auch auf Daten vor 1987 anwendbar sei. Begründet wurde dies damit, dass bei mit Auskunftspflicht erhobenen Daten für die Datenweitergabe alleine der Zeitpunkt der Datenanfrage entscheidend sei. Zudem würde ein

schutzwürdiges Vertrauen der Befragten, welches durch eine rückwirkende Anwendung des BStatG 87 verletzt sein könnte nicht vorliegen, da die Befragten aufgrund der Auskunftspflicht keine Wahl gehabt hätten.

## 2.2 Datenauswahl

Mit dieser überraschenden Wende in der Interpretation des BStatG durch das FDZ Bund war der Weg frei für die nächsten konkreten Arbeitsschritte. Da nicht alle Mikrozensus zwischen 1957 und 1987 erworben werden können, wurde in einem Arbeitspapier des German Microdata Lab begründet, welche Daten für die Sozialwissenschaften von besonderem Interesse sind und deshalb als Scientific Use File für die Wissenschaft bereitgestellt werden sollten (im Detail Lüttinger/Wirth 2004: Auswahl relevanter Datenzeitpunkte für die Bereitstellung von Scientific Use Files von Mikrozensus vor 1989. German Microdata Lab Arbeitspapier 2004/02). Die konkrete Festlegung der Zeitpunkte orientierte sich an folgenden Kriterien:

Von primärer Bedeutung sollte die Relevanz der für die einzelnen Erhebungszeitpunkte verfügbaren Informationen für sozial- und wirtschaftswissenschaftliche Fragestellungen sein. Zudem galt es zu berücksichtigen, ob die relevanten Merkmale z.B. im Grundprogramm jährlich oder nur in größeren Abständen (Zusatz- und Ergänzungsprogramme ab 75) erfragt werden (z.B. Ausländerangaben; Ausbildung) bzw. nur über Zusatzbefragungen (bis 1975) erhoben wurden (z.B. Bildungsangaben).

Darüber hinaus spielt ebenfalls eine Rolle, dass die Anzahl und Art der erhobenen Merkmale in den einzelnen Erhebungen nicht von Jahr zu Jahr schwanken, sondern in der Regel für längere Zeiträume durch Mikrozensusverordnungen bzw. Gesetze festgelegt sind. Für den Zeitraum zwischen 1957 und 1989 sind dies vier Phasen: 1957-1961, 1962-1974, 1975-1982 und 1985-1990, innerhalb derer die Erhebungsprogramme nahezu unverändert durchgeführt wurden. Somit besteht bei der Auswahl von konkreten Zeitpunkten ein gewisser Spielraum, d.h. es ist aus einer „inhaltlichen“ Perspektive weitgehend unerheblich, ob z.B. der Mikrozensus 1964 oder der Mikrozensus von 1969 bereitgestellt wird, da sich das Frageprogramm zwischen 1962 und 1974 kaum geändert hat.

Zwischen 1962 und 1974 wurden aufgrund des relativ begrenzten Programms der Grunderhebung und zur Erhöhung der Flexibilität des Fragenprogramms zahlreiche Zusatzerhebungen zum Mikrozensus durchgeführt. Die Zusatzerhebungen sind auf spezifische Fragestellungen konzentriert, die das jeweilige Grundprogramm ergänzen. Folgende Zusatzprogramme waren u.E. wichtige Datenquellen für die Forschung:

- a) Zusatzerhebung 1964, 69, 70 mit dem Themenschwerpunkt „Berufliche Ausbildung“;
- b) Zusatzerhebung April 1971 „Berufliche und soziale Umschichtung der Bevölkerung“;

- c) Zusatzerhebung Juli 1972 mit dem Themenschwerpunkt „Ausbildungsabsichten der Eltern für ihre Kinder“;
- d) Die Zusatzerhebung April 1974 (0,1 %) mit dem Themenschwerpunkt „Berufliche Mobilität von Frauen“.

Ein letztes Kriterium für die Auswahl konkreter Zeitpunkte war schließlich, dass möglichst eine kontinuierliche Zeitreihe anzustreben war. So empfiehlt das KVI Gutachten „eine Datenserie in einem Abstand von fünf Jahren für die Reihe bis Ende der siebziger Jahre und im Abstand von zwei Jahren für die aktuelleren Daten“ (KVI Gutachten: 295). Unter Zugrundelegung der oben genannten Kriterien ergaben sich folgende Zeitpunkte:

1957, 1964 (Zusatzerhebung), 1969 (Zusatzerhebung), 1970 (Zusatzerhebung), 1971 (Zusatzerhebung), 1972 (Zusatzerhebung), 1973, 1974 (Zusatzerhebung), 1976, 1978, 1980, 1982, 1985 und 1987.

Die Recherchen durch das Forschungsdatenzentrum des Statistischen Bundesamtes ergaben, dass die Daten aus den fünfziger und sechziger Jahren sowie die Daten der Zusatzerhebungen nicht mehr verfügbar waren, sondern erst ab dem Mikrozensus 1973 noch Daten und entsprechende Metainformationen vorlagen. Daraufhin wurde beschlossen, in einem ersten Arbeitsschritt für die Mikrozensen 1973, 1976 und 1982 Scientific Use Files zu erstellen. Darüber hinaus beantragte ZUMA im Februar 2005 für vier weitere Mikrozensen, die Erhebungsjahre 1978, 1980, 1985 und 1987, die Erstellung von Scientific Use Files.

### 2.3 Erstellung und Prüfung der Scientific Use Files

In Gesprächen mit dem FDZ Bund wurden ab Mitte 2004 verschiedene Optionen diskutiert, die Scientific Use Files der älteren Mikrozensen zu erstellen. Im Ergebnis war vorgesehen, dass ZUMA die Ziehung der Stichproben übernimmt und zweckgebunden die Daten hierfür zur Verfügung gestellt bekommt. Dieses Vorgehen wurde jedoch nicht realisiert, da das FDZ der Länder diesem Verfahren nicht zustimmte und im Dezember 2004 die Stichprobenziehung und die Erstellung der Scientific Use Files übernahm. ZUMA sollte die Scientific Use Files der Mikrozensen 1973, 1976, 1982 im Februar des Jahres 2005 vom FDZ der Länder zur Prüfung erhalten. Aus verschiedenen Gründen wurde dieser Termin vom FDZ nicht eingehalten. Um den Prozess der Erstellung der Scientific Use Files zu beschleunigen bot das GML daraufhin dem Forschungsdatenzentrum der Statistischen Landesämter an, in Zusammenarbeit mit dem FDZ für jeden Datensatz Vorschläge zur Anonymisierung der Daten zu unterbreiten. Diese Arbeiten wurden im Sommer 2005 umgesetzt. Die Scientific Use Files der Mikrozensen 73, 76 und 82 wurden schließlich Anfang November 2005 dem GML zur Erzeugung der Setups, der Erstellung von Datendokumentationen und zur Plausibilitäts-



prüfung übermittelt. Die Plausibilitätsprüfung der Mikrozensus Scientific Use Files erfolgt durch den Vergleich des aufbereiteten Datensatzes mit den vom Statistischen Bundesamt veröffentlichten Tabellen in den entsprechenden Reihen der Fachserie. Hierbei muss nach den vom Statistischen Bundesamt verwendeten Bevölkerungs- und Erwerbskonzepten gewichtet und selektiert werden. Die Ergebnisse dieser Plausibilitätsprüfung werden dem Nutzer über das WWW zur Verfügung gestellt, sodass er z.B. anhand dieser Zahlen seine eigenen, mit dem Scientific Use File erzeugten Zahlen vergleichen kann.<sup>1</sup> Diese Arbeiten waren Ende Januar 2006 abgeschlossen und die Ergebnisse der Arbeiten dem FDZ übermittelt. Im Oktober 2006 standen die Scientific Use Files über das FDZ-L (NRW) für die Forschung zur Verfügung.

## 2.4 Wissenschaftlicher Service

Der wissenschaftliche Service im Projekt „Rückerschließung“ umfasste die Erstellung standardisierter Datendokumentationen, die Erstellung von Mikrodaten-Tools sowie die Betreuung von Gastwissenschaftlern während ihrer Forschungsaufenthalte beim GML.

### **Erstellung standardisierter Dokumentationen**

Für die optimale Nutzung der Daten der Scientific Use File der Mikrozensen 1973, 1976 und 1982 wurden ausführliche Informationen und Dokumentationen zu den aufbereiteten Daten für den Nutzer erstellt, die alle über das Web-Angebot des GML abrufbar sind. Diese Informationen umfassen Aspekte der Datenaufbereitung, gesetzliche Grundlagen, Erhebungsmodalitäten, Klassifikationen, thematische Übersichten etc. (siehe auch Übersicht 1, Kap. 2.1). Zudem werden Setups erstellt und über das WWW zum Download angeboten, mit denen die im ASCII Format gelieferten Rohdaten eingelesen werden können. Diese Setups werden für die häufig benutzten Programme SPSS, Stata und SAS bereitgestellt.

### **Methodenforschung/Mikrodatentools**

Neben den Serviceleistungen und der Wissensvermittlung werden im GML eigene Forschungsarbeiten durchgeführt. Ein Schwerpunkt im Projekt „Rückerschließung“ lag in der Entwicklung von Analysewerkzeugen, mit denen die Nutzer auf Grundlage amtlicher Daten sozialwissenschaftlich interessante Konzepte und Klassifikationen umsetzen können. Die Bereitstellung solcher Tools hat viele Vorteile. Insbesondere entfällt diese aufwändige Arbeit für den Nutzer und es wird eine Standardisierung erreicht, welche die Validität und Replizierbarkeit der erzielten Befunde verbessert.

---

<sup>1</sup> Siehe z.B.: [www.gesis.org/Dauerbeobachtung/GML/Daten/MZ/mz\\_1976/randverteilungen\\_mz76.htm](http://www.gesis.org/Dauerbeobachtung/GML/Daten/MZ/mz_1976/randverteilungen_mz76.htm)

Unser Schwerpunkt im Berichtszeitraum lag in der Umsetzung zweier Bildungsskalen, die allen Interessierten uneingeschränkt über das Internet zur Verfügung gestellt wurden.<sup>2</sup> Es handelt sich dabei um die Umsetzung der Bildungsskala ISCED-1997 sowie der CASMIN-Bildungsskala (siehe Kap. 2.3.1). Beide Bildungsskalen wurden für die gegenwärtig verfügbaren amtlichen Mikrodaten erstellt. (Volkszählung 1970, Mikrozensus-Zusatzerhebung 1971, Mikrozensus 1976-2004).

### **Betreuung von Gastwissenschaftlern**

Das GML bietet interessierten Wissenschaftlern die Möglichkeit, die Daten des Mikrozensus sowie weitere Mikrodaten der amtlichen Statistik im GML zu analysieren. Im Projekt „Rückerschließung“ wurden im Berichtszeitraum 5 Gastwissenschaftler (mit z.T. mehreren Aufenthalten) betreut, die mit älteren Daten der amtlichen Statistik gearbeitet haben (siehe Kap. 2.2.2).

## **2.5 Daten der Volks- und Berufszählung 1970**

Die Volks- und Berufszählung von 1970<sup>3</sup> wurde in zwei Teilen durchgeführt. Als Vollerhebung wurden für die gesamte Bevölkerung Merkmale erfragt, die für regional und fachlich tief gegliederte Auswertungen nötig waren. Einer 10 Prozent Stichprobe der Bevölkerung wurden zusätzliche Fragen gestellt, die weitergehende demographische, wirtschaftliche und soziale Strukturanalysen ermöglichen sollten. Diese 10 Prozent Stichprobe der Volkszählung 1970 wurde in anonymisierter Form vom Projekt „Vergleichende Analysen der Sozialstruktur mit Massendaten“ (VASMA)<sup>4</sup> der Universität Mannheim Ende der 1970er Jahre von den Statistischen Landesämtern erworben. Im VASMA Projekt wurde aus rechentechnischen Gründen daraus eine 1 Prozent Stichprobe für wissenschaftliche Analysen gezogen.

Das VASMA-Projekt wurde von der Stiftung Volkswagenwerk auch unter dem Gesichtspunkt gefördert, langfristig die Bildung einer Institution zu fördern, deren Aufgabe es sein sollte, für die Sozialwissenschaften die Daten der amtlichen Statistik in einer für die Forschung adäquaten Weise zugänglich zu machen. Nach vielfältigen Bemühungen wurde 1987 bei ZUMA die Abteilung Mikrodaten (seit 2003: German Microdata Lab) eingerichtet, welche die im VASMA-Projekt begonnenen Arbeiten fortsetzte und seitdem die Daten der Volkszählung von 1970 pflegte.

Diese Daten stellen eine einzigartige und unersetzliche Datenquelle für die Forschung dar, auch deshalb, weil diese Daten bei den statistischen Ämtern nicht mehr vorlagen. Daher

---

<sup>2</sup> <http://www.gesis.org/Dauerbeobachtung/GML/Service/Mikrodaten-Tools/index.htm>

<sup>3</sup> Zum Inhalt der Daten siehe: [www.gesis.org/Dauerbeobachtung/GML/Daten/VZ1970/index.htm](http://www.gesis.org/Dauerbeobachtung/GML/Daten/VZ1970/index.htm)

<sup>4</sup> [www.sowi.uni-mannheim.de/lehrestuehle/lesas/forschung/vasma/vasma.html](http://www.sowi.uni-mannheim.de/lehrestuehle/lesas/forschung/vasma/vasma.html)

wurden die 10 Prozent Originaldaten der VZ70 im Rückerschließungsprojekt aufbereitet<sup>5</sup> und dem Forschungsdatenzentrum der Statistischen Landesämter (NRW) übermittelt, so dass nun auch die amtliche Statistik wieder über diese Daten verfügt.

Das GML sah zudem vor, die im VASMA-Projekt gezogene 1 Prozent Stichprobe der VZ70 direkt über das ZA der Wissenschaft zur Verfügung zu stellen, da die Volkszählungsdaten nach Ansicht des GML (sowohl die 1 % Stichprobe, als auch die 10 % Stichprobe) als absolut anonym (Public Use File) zu betrachten sind und daher nicht unter die Bestimmungen des §16, BStatG 1987. Begründet wurde dies mit folgenden Punkten:

- Schon die damalige Übermittlung der Daten an das VASMA-Projekt erfolgte in anonymisierter Form. Zu diesem Zweck hatten die Landesämter eine Reihe von Merkmalen (Variablen 2, 3 und 186-205) gelöscht. Darüber hinaus wurden keine kleinräumigen Regionalinformationen (Regierungsbezirk, Kreis, Gemeinde) weitergegeben.
- Weiterhin wurde nur die Substichprobe von 10 % an das VASMA-Projekt weitergegeben. Aus dieser Substichprobe wurde im Rahmen des VASMA-Projektes wiederum eine 10 %-Substichprobe gezogen wurde (=> VZ70 1 %-Substichprobe). Dies entspricht den Empfehlungen des Anonymisierungsprojekts für die Generierung eines CAMPUS-Files, wenngleich es damals nicht zu diesem Zweck gedacht war.
- Unabhängig hiervon spielt das Alter der Daten eine ganz zentrale Rolle für die Einstufung der Daten (10 % und 1 %) als absolut anonym. Die Datenerhebung liegt inzwischen mehr als 30 Jahre zurück. Dies bedeutet mit Blick auf die relevante Population:
  - Mit hoher Wahrscheinlichkeit sind nahezu alle damals über 55jährigen inzwischen verstorben (d.h. Personen, die gegenwärtig 90 Jahre und älter sind);
  - die heute 0-35jährigen sind in der Stichprobe von 1970 nicht enthalten;
  - für einen Großteil der heute 35-55jährigen gibt es für 1970 zu vielen Informationen keine Angaben (z.B. Angaben zu schulischer, beruflicher Ausbildung, Erwerbstätigkeit, Familie etc.);
  - hinzu kommen massive Migrationsprozesse (d.h. nur noch ein geringer Teil der Haushalte dürfte überhaupt in den damaligen Auswahlbezirken - zu denen keine Informationen im Datensatz vorliegen - wohnen).
- Mit Blick auf das Zusatzwissen wurde auf Südfeld (1987) verwiesen, dessen Aussage sich auf die absolute Anonymisierung von Daten bezieht: „Je älter die Einzelangaben werden, desto geringer wird einerseits das Reidentifikationsrisiko, weil die Verfügbarkeit kompatiblen Zusatzwissens immer unwahrscheinlicher wird und das Interesse eines

---

<sup>5</sup> Lechert, Yvonne und Paul Lüttinger, 2006: Volks- und Berufszählung 1970 (10 % Stichprobe): Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/11.

potentiellen Angreifers geringer wird.” (Südfeld 1987: 148<sup>6</sup>). Für die Daten der Volkszählung von 1970 existiert kein bekanntes Zusatzwissen, mit dem eine Deanonymisierung erfolgen könnte.

Das Forschungsdatenzentrum der Statistischen Landesämter (NRW) folgte dieser Begründung bislang jedoch nicht, sodass die Auslieferung der Volkszählungsdaten über das ZA bis zur endgültigen Klärung der Frage nach dem Grad der Anonymität der Volkszählungsdaten, aufgeschoben wurde.

## 2.6 Nutzen und Verwertbarkeit

Das langfristige Ziel des Projektes Rückerschließung bestand in der grundsätzlichen Klärung der Frage, ob die Mikrozensushebungen vor 1989 in Bezug auf die Weitergabe der Daten an die Wissenschaft ebenfalls dem Bundesstatistikgesetz von 1987 unterliegen, das eine Form der Weitergabe amtlicher Mikrodaten in faktisch anonymer Form erlaubt oder ob die bis zum Projektbeginn praktizierte Form der ausschließlichen Weitergabe als absolut anonymisierte Daten auch zukünftig beizubehalten ist. Der Nutzen des Projektes bestand primär in der Lösung dieser Frage zugunsten der Weitergabe in Form von Scientific Use Files sowie in der unmittelbaren Verwertung dieses Projektergebnisses, indem drei Scientific Use Files von Mikrozensus vor 1989 der Wissenschaft zur Verfügung gestellt wurde. Die über das Projekt hinaus gehende Verwertbarkeit besteht darin, dass generell die Gültigkeit des BStatG 1987 von Seiten der amtlichen Statistik in Bezug auf die Weitergabe amtlicher Daten anerkannt wurde.

## 2.7 Fortschritt bei anderen Stellen

Neben dem Rückerschließungsprojekt, das in enger Kooperation mit den Forschungsdatenzentren des Statistischen Bundesamtes und der Statistischen Landesämter durchgeführt wurde, haben sich keine anderen Stellen mit dem Problem der Weitergabe von Mikrozensus vor 1989 als Scientific Use File beschäftigt.

---

<sup>6</sup> Südfeld, E., 1987: Anonymisierungsstandards und generelle Abwicklungsregelungen für Anforderungen nach anonymisierten Einzelangaben im Statistischen Bundesamt. S. 146-157 in: Statistisches Bundesamt 1987 (Hrsg.).

## 2.8 Erfolgte/geplante Veröffentlichung des Ergebnisses

Neben den über das Web verfügbaren zahlreichen Dokumentationen und Informationen zu den Scientific Use Files 1973, 1976 und 1982, die im Projekt Rückerschließung erarbeitet wurden, gibt es folgende Veröffentlichungen aus dem Projekt:

- Veröffentlichungen in Zeitschriften

Lüttinger, Paul; Wirth, Heike Lüttinger, Paul und Heike Wirth, 2004: Zur Weitergabe von älteren amtlichen Mikrodaten an die Wissenschaft. Allgemeines Statistisches Archiv 88 (4): 473-486.

- Vorstellung des Projektes

Lüttinger, Paul, Bernhard Schimpl-Neimanns, Heike Wirth und Georgios Papastefanou, 2004: The German Microdata Lab at ZUMA: Services Provided to the Scientific Community, S. 455-467 in: Schmollers Jahrbuch („Zeitschrift für Wirtschafts- und Sozialwissenschaften“), 3/2004, Berlin: Duncker & Humblot.

Lüttinger, Paul, Bernhard Schimpl-Neimanns, Heike Wirth und Georgios Papastefanou, 2003: Mikrodaten (German Microdata Lab): Das Servicezentrum für amtliche Mikrodaten bei ZUMA. ZUMA-Nachrichten 52:153-172.

Lüttinger, Paul, 2006: Mikrozensus 1973, 1976 und 1982 als Scientific Use Files verfügbar. ZUMA. ZUMA-Nachrichten 59:137-138.

- Arbeitspapiere/Berichte des Projektes

Lüttinger, Paul und Heike Wirth, 2004: Regelungen zur Weitergabe von amtlichen Mikrodaten an die Wissenschaft in den Bundesstatistikgesetzen seit 1953. Arbeitspapier German Microdata Lab, Arbeitspapier 2004/01.

Lüttinger, Paul und Heike Wirth, 2004: Auswahl relevanter Datenzeitpunkte für die Bereitstellung von Scientific Use Files von Mikrozensus vor 1989. German Microdata Lab Arbeitspapier 2004/02.

Lechert, Yvonne, Julia H Schroedter und Paul Lüttinger, 2006: Die Umsetzung der Bildungsklassifikation CASMIN für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensus 1976-2004. ZUMA-Methodenbericht 2006/12.

Lechert, Yvonne und Paul Lüttinger, 2006: Volks- und Berufszählung (10 % Stichprobe) Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/11.

Lechert, Yvonne, Paul Lüttinger und Schunck, Reinhard, 2006: Vergleich der Mikrozensus Scientific Use Files 1976 und 1982. ZUMA-Methodenbericht 2006/10.

Lechert, Yvonne, Paul Lüttinger und Schunck, Reinhard, 2006: Vergleich der Mikrozensus Scientific Use Files 1973 und 1976. ZUMA-Methodenbericht 2006/09.

- Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1982. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/03.
- Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1976. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/02.
- Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1973. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/01.
- Schroedter, Julia H., Yvonne Lechert und Paul Lüttinger, 2006: Die Umsetzung der Bildungsskala ISCED-1997 für die Volkszählung 1970, die Mikrozensus- Zusatzerhebung 1971 und die Mikrozensen 1976-2004. ZUMA-Methodenbericht 2006/08.
- Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1973 Codebuch und Grundauszählung: ZUMA-Interner Projektbericht 04/2006.
- Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1976 Codebuch und Grundauszählung: ZUMA-Interner Projektbericht 05/2006.
- Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1982 Codebuch und Grundauszählung: ZUMA-Interner Projektbericht 06/2006.

### 3 Mikrodaten-Informationssystem MISSY

Das Projekt MISSY (Mikrodaten-Informationssystem)<sup>7</sup> hatte zum Ziel, der Scientific Community datenrelevante Informationen – so genannte Metadaten – in systematischer Weise zur Verfügung zu stellen.

Für eine Nutzung von amtlichen Mikrodaten seitens der Wissenschaft ist die Aufbereitung und Bereitstellung von Metadaten unabdingbar. Nur mit dem Wissen über die Eigenschaften des Datensatzes und seiner einzelnen Variablen können die Daten von Forschern sachgerecht genutzt und angemessen interpretiert werden. Somit lag ein Schwerpunkt des Projektes auf der Bereitstellung dieser Metadaten. Da das Projekt Pilotcharakter aufweist, wurden exemplarisch zwei Jahrgänge des Mikrozensus ausgewählt, für die die Metadaten zum Teil neu erstellt, aufbereitet und bereitgestellt wurden. Die Bereitstellung oder das Erstellen von ‚Wissen‘ bzw. Informationen allein reichen jedoch nicht aus, um Forschern die Beschaffung von datenrelevanten Informationen zu erleichtern. Gerade mit der Entwicklung der Kommunikations- und Informationstechnologie und damit einer wachsenden Menge an verfügbarem Wissen ist nicht unbedingt mehr das Wissen selbst, sondern das Wissen darüber, in welcher Weise dieses Wissen zugänglich ist, entscheidend für dessen Nutzung. Der zweite Schwerpunkt von MISSY bestand dementsprechend in der Aufgabe, den Zugang zu den Metadaten möglichst effektiv zu gestalten, so dass Forscher einfach und zeitsparend die gewünschten Informationen herausfiltern und abrufen können. Dazu gehört neben der systematischen Organisation der Metadaten auch die Verknüpfung von Informationen sowohl innerhalb einzelner Datensätze als auch – zum Zwecke der Vergleichbarkeit der Erhebungsjahre – über verschiedene Datensätze hinweg.

Bei MISSY können zwei Aspekte unterschieden werden: Zunächst der inhaltliche Aspekt, also die Frage, *welche* Informationen das System bereitstellt. Hier kann zwischen variablenbezogenen Informationen und studienbezogenen Informationen unterschieden werden. Der zweite Aspekt bezieht sich auf die Organisation und die Struktur des Informationssystems, also *in welcher Weise* auf die Informationen zugegriffen werden kann. Dabei spielen neben dem strukturellen Aufbau des Systems vor allem die Verknüpfungen zwischen einzelnen Informationen eine große Rolle. Im Folgenden werden zunächst die in MISSY eingebundenen Variableninformationen und die verfügbaren Zugänge zu diesen Informationen beschrieben. Anschließend werden Umfang und Inhalt der studienbezogenen Informationen erläutert. Im vierten Abschnitt wird auf eine Usability-Studie hingewiesen, die 2006

---

<sup>7</sup>Das System ist unter <http://www.gesis.org/Dauerbeobachtung/GML/MISSY/> frei verfügbar

zur Optimierung des Informationssystems durchgeführt wurde. Die technische Umsetzung ist Thema des fünften Abschnitts. Abschließend wird noch auf einige Rahmendaten hingewiesen.

Zum generellen Aufbau von MISSY: Die Eingangsseite von MISSY informiert über Ziel und Inhalt des Informationssystems und gibt einen ersten Überblick über den Mikrozensus. Detailliertere Informationen sowohl zu MISSY als auch zum Mikrozensus sind über diese Seite abrufbar. Über eine links platzierte Navigationsleiste erhält der Nutzer Zugang zu variablen- oder studienbezogenen Informationen, die zur besseren Differenzierbarkeit unterschiedlich farblich gekennzeichnet sind. Diese farbliche Differenzierung wird auf allen Seiten in den Überschriften übernommen, was zu einer einfachen Orientierung im System beitragen soll. Die Navigationsleiste bleibt auf allen Seiten bestehen und ist somit von jedem Punkt in MISSY aus anwählbar. Ebenfalls zu jedem Zeitpunkt ist auf die Suchfunktion, die Eingangsseite von MISSY oder auf die Internetauftritte der GESIS bzw. des German Microdata Labs zugreifbar.



### 3.1 Variablenbezogene Informationen

Die Informationen über eine Variable werden in MISSY auf je einer Seite zusammengefasst, so dass alle zum Arbeiten mit dieser Variablen notwendigen Informationen entweder auf dieser Seite zu finden sind oder von dieser Seite aus verlinkt werden. Folgende Abbildung bietet einen Überblick über Umfang und Darstellung der Variableninformationen am Beispiel der Variable zur geringfügigen Beschäftigung vom Mikrozensus 1997:

Abbildung 3: Variableninformationen für das EF99

Mikrodaten-Informationssystem

Mikrozensus Grundfile **1997** Variable **EF99**

HOME [GESIS](#) [GML](#)

[<](#) [>](#)

---

**F24a Geringfüg. Beschäftigung (Berichtsw.): einzige oder hauptsächl. Erwerbstätig.**

■ **Thematische Gliederung:**

■ **Andere Erhebungszeitpunkte für diese Variable:**

**Fragebogen:**

**Fragenummer und -text:**

**Erläuterungen zur Frage im Anhang:**

**Filteranweisung:**

**Filterangaben:**

**Filterangaben (formal):**

**Substichprobe:**

**Auswahlsatz:**

**Auskunftspflicht:**

■ **Weitere Informationen zu dieser Variable:**

Arbeitsmarkt und Erwerbsbeteiligung > Erwerbsstatus

2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1993	1991	1989
EF99	EF99	EF99	EF99	EF99	EF99	EF99	EF99	EF99	EF75	EF75	EF75	EF75

Die Variable ist für die Mikrozensus 1989-1995 nicht verfügbar, es besteht aber Vergleichbarkeit mit anderen Variablen. Diese sind in der Tabelle rot markiert.

Erhebungsbogen 1+E

24a: Handelte es sich bei dieser geringfügigen Beschäftigung um Ihre **einzige** oder Ihre **hauptsächliche** Erwerbs- oder Berufstätigkeit?

-

Anweisung 20: Für Personen im Alter von 15 Jahren und älter

Befragter hat in der Berichtswoche eine geringfügige Beschäftigung ausgeübt.

IF EF98 == 1

-

1,00%

ja

[Schlüsselverzeichnis](#) [Fragebogen](#) [Interviewerhandbuch](#)

**Häufigkeitsauszählung**

Value Label	Value	Frequency	%	Valid %
Ja	1	12 105	2,4%	88,0%
Nein	8	1 973	0,4%	14,0%
Valid Total		14 078	2,8%	100,0%
(M) Entfällt	0	495 814	97,2%	
Total		509 892	100,0%	

Datenbasis: Mikrozensus Scientific Use File 1997

Auf diesen Seiten werden umfassende Informationen zu den einzelnen Variablen zusammengefasst:

- Zunächst werden Jahr und Variablenname angezeigt, die daneben platzierten Pfeile ermöglichen es, innerhalb der Variablenliste (d.h. nach aufsteigender oder absteigender EF-Nummer) zu blättern.
- Die *Variablenlabels* wurden sprachlich überarbeitet und vereinheitlicht und folgen einem bestimmten Schema. Diese sprachliche Überarbeitung stellte einen wichtigen Schwerpunkt des MISSY-Projekts dar (Granato 2001: 41ff), da nur mit einheitlichen Variablenlabels eine intertemporale Vergleichbarkeit der Variablen hergestellt werden kann. Die im

Rahmen dieses Projekts erarbeiteten Kriterien zur Variablenlabelbildung bilden mittlerweile auch die Grundlage für die Aufbereitung der neuen Scientific Use Files.

- Darüber hinaus gibt es Angaben über die *thematische Zugehörigkeit* der Variablen mit der Option, auf die einzelnen Ebenen der thematischen Gliederung zu wechseln (vgl. 3.2).
- Eine weitere Übersicht informiert über die *Erhebung des Merkmals zu weiteren Zeitpunkten*. Hier können die Variablen für andere Erhebungszeitpunkte direkt angeklickt werden; zudem werden auf dieser Ebene auch Angaben über die Vergleichbarkeit von Merkmalen über die Zeit abgelegt (vgl. Zugänge Matrix, Abb. 5). Mit Hilfe von roten Markierungen werden Variablenbrüche gekennzeichnet, über einen Informationsbutton können dann die Informationen zur Vergleichbarkeit abgerufen werden. Kann eine Vergleichbarkeit vor und nach 1996 hergestellt werden, wird von dieser Stelle aus auf eine Seite der Studieninformationen verlinkt, auf der SPSS-Routinen abrufbar sind.
- *Fragennummer- und text, Name des Fragebogens*: Die Angaben und Erläuterungen zur Frage und auch zur Filterführung des jeweiligen Merkmals (s.u.) wurden mit dem Microsoft-Programm InfoPath (vgl. 3.5) in das System eingespeist. Die Informationseingabe erfolgte dabei manuell anhand schriftlicher Vorlagen.
- *Erläuterungen im Fragebogen*: Hier werden zusätzliche Informationen zu den Fragen aus dem Fragebogen dokumentiert; auch diese Aufbereitung erfolgte manuell.
- *Häufigkeitsauszählung*: Die Häufigkeitsauszählungen wurden aus dem Setup generiert und umfassen neben der Valuelabels eine ungewichtete Häufigkeitsauszählung und die Angaben der allgemeinen und der gültigen Prozentangaben. Die Reihenfolge der Darstellung ist dabei modifizierbar. Bei den wenigen Variablen, die über sehr viele Merkmale verfügen wie z.B. bei den Angaben zu Wirtschaftszweigen oder der Klassifikation der Berufe, werden zunächst Zusammenfassungen der Merkmale präsentiert, und die Vollversion ist über eine Auswahl aufrufbar.
- *Filterführungen*: Die Filterführungen im Fragebogen werden sowohl sprachlich als auch formal für alle Filterstufen nachvollzogen. Zusätzlich werden die bei der Filterung relevanten Variablen verlinkt, was die Nachvollziehbarkeit der Angaben erhöht. Diese Informationen existierten bislang nur indirekt durch den Fragebogen und wurden für MISSY erstellt.
- *Filteranweisungen*: Neben den Filterführungen, die im Fragebogen mit Pfeilen gekennzeichnet sind, gibt es im Mikrozensus schriftliche Filteranweisungen, die ab 1996 mit Fragennummern versehen sind. Bei den Anweisungen beschränkt sich die Darstellung auf die sprachliche Dokumentation, da sie sich formal nicht immer ausdrücken lassen. Auch diese Informationen wurden manuell erstellt.
- *Auswahlsatz und Auskunftspflicht*: Es wird dokumentiert, ob die Frage zu einer Substichprobe gehört und deshalb nur von einem bestimmten Prozentsatz der Befragten zu beantworten ist und ob die Frage für alle oder einen Teil der Befragten auf freiwilliger Basis beantwortet wird.
- *Generierung (nicht abgebildet)*: Bei den Variablen, die auf Basis anderer Merkmale gebildet wurden, wird diese Bildung nachvollzogen und erklärt. Diese Informationen wurden anhand von Unterlagen des Statistischen Bundesamtes erstellt und sind zusätzlich über die Studieninformationen abrufbar.
- *Schlüsselverzeichnis, Fragebogen, Interviewerhandbuch*: Mit den Verlinkungen zu den relevanten Seiten des Schlüsselverzeichnis, Fragebogens und Interviewerhandbuchs werden Verbindungen zwischen variablen- und studienbezogenen Informationen angeboten, die die Such- und Arbeitszeiten erheblich verringern können. Die Markierung innerhalb

der pdf-Dokumente erfolgte mittels Sprungadressen, so dass die Dokumente einzeln editiert wurden.

- *Definitionen (nicht abgebildet)*: Die über die Studieninformationen zugänglichen Informationen über Konzepte und Definitionen, auf deren Basis der Mikrozensus umgesetzt wird, werden zusätzlich mit den Variableninformationen der Variablen verlinkt, für deren Themenbereich es einen Eintrag im Definitionskatalog gibt. Die Verlinkung führt dabei exakt zur relevanten Stelle im Katalog, so dass Suchzeiten verringert werden können.

### 3.2 Zugänge zu variablenbezogenen Informationen

Neben den Verbindungen zwischen den variablen- und studienbezogenen Informationen liegen die Vorteile der Datenbank im *Navigationssystem*, das einen vielfältigen und bedürfnisangepassten Zugang zu den Variableninformationen erlaubt. Dabei gehen wir davon aus, dass sich unterschiedliche Nutzergruppen auch in ihrem Informationsstand über den Mikrozensus unterscheiden und deshalb ein unterschiedliches Maß an Hilfestellung durch das Informationssystem benötigen. Zudem unterscheiden sich Suchstrategien auch nach verschiedenen Zielsetzungen. In MISSY werden vier Zugangswege zu den Variableninformationen angeboten:

#### a) Zugang über Thematische Gliederung

Diese Zugangsmöglichkeit ist insbesondere dafür geeignet, sich einen ersten Überblick über die im Mikrozensus enthaltenen Themenschwerpunkte zu verschaffen oder sich über die vorhandenen Merkmale zu einem speziellen Thema zu informieren. Die thematische Gliederung umfasst auf der ersten Ebene elf Kategorien, die in ihrer Bezeichnung an den Thesaurus des IZ angeglichen und somit für entsprechende Suchmaschinen kompatibel sind. Auf der zweiten Stufe der Kategorien befinden sich Links, die auf die Literaturliste verweisen. Dabei wird aber nicht die gesamte Literaturliste aufgerufen, sondern es werden in alphabetischer Reihenfolge nur die Literaturangaben ausgegeben, die diesem Thema zuzuordnen sind. Somit erhält ein Nutzer, der z.B. erste Recherchen zu einem thematischen Schwerpunkt des Mikrozensus unternimmt, früh Hinweise auf bereits veröffentlichte Forschungsergebnisse zu dem interessierenden Thema. Weiterhin werden innerhalb dieser thematischen Gliederung Verknüpfungen zu Forschungsergebnissen in der Rubrik Tabellen angeboten (vgl. 3.7). Die folgende Grafik dokumentiert den Zugang über die thematische Gliederung:

Abbildung 4: Zugang über thematische Gliederung

**Mikrodaten-Informationssystem**

HOME GESIS GML

**MISSY**

**Thematische Gliederung**

**Variableninformationen**

- Thematische Gliederung
- Variablenliste
- Variablen-Zeitpunkte-Matrix

**Allgemeine Informationen**

- Erhebung
- Daten
- Klassifikationen
- Literatur

Suche

Links


FAQ

Impressum

- Demographie und Bevölkerung
  - Daten zur Person
    - Alter
    - Geschlecht
    - Familienstand
      - Familienstand
      - Familienstand: Haushaltsbezugsp.
      - Familienstand: Familienbezugsp.
      - Familienstand: Lebenspartner der Haushaltsbezugsp.
    - Eheschließungsjahr
    - Staatsangehörigkeit
  - Bevölkerungstyp
  - Regionalangaben
- Ausländer und Migration
- Arbeitsmarkt und Erwerbsbeteiligung
- Unterhalt und Einkommen
- Sozialversicherung und Vorsorge
- Bildung und Qualifikation
- Pendler
- Privathaushalt und Familie
  - Haushaltsstruktur
  - Familienstruktur
  - Struktur von Lebensgemeinschaften
- Wohnverhältnisse
- Gesundheit
- Stichprobe

Kategorien verzweigen sich – je nach Anzahl und thematischem Umfang der Variablen – unterschiedlich stark in Unterkategorien. Die letzte Stufe (rot markiert wie in der Abbildung bei der Kategorie „Ausländer und Migration“ oder „Familienstand“ zu sehen) enthält eine Liste der in allen Erhebungen zu dem jeweiligen Thema verfügbaren Merkmale. Nach der Auswahl eines Merkmals werden in einem weiteren Schritt sämtliche Erhebungszeitpunkte aufgelistet, für die das Merkmal in den Scientific Use Files verfügbar ist, so dass der Nutzer sich den gewünschten Erhebungszeitpunkt anzeigen lassen kann. Dabei werden die im Mikrozensus ab 1996 umgesetzten Veränderungen in der Erhebung und damit verbundenen Variablenbrüche auf der letzten Ebene der thematischen Gliederung dokumentiert. Die roten Markierungen geben die nicht identischen, aber vergleichbaren Variablen ab 1996 an. Für die Erhebungszeitpunkte 1995 und 1997 sind die Zugänge zu den Variableninformationen umgesetzt:

Abbildung 5: Variablenbrüche auf Ebene der thematischen Gliederung



**Variableninformationen**

- Thematische Gliederung
- Variablenliste
- Variablen-Zeitpunkte-Matrix

**Allgemeine Informationen**

- Erhebung
- Daten
- Klassifikationen
- Literatur

Suche  >

- Links
- FAQ
- Impressum

**Mikrodaten-Informationssystem**

**Mikrozensus: Erhebungszeitpunkte**


**Differenz zur normalen Arbeitszeit (Berichtsw.): Grund**

2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1993	1991	1989
EF144	EF144	EF144	EF144	EF144	EF144	EF144	EF144	EF144	EF144	EF101	EF101	EF101
EF145	EF145	EF145	EF145	EF145	EF145	EF145	EF145	EF145	EF145	EF101	EF101	EF101
EF146	EF146	EF146	EF146	EF146	EF146	EF146	EF146	EF146	EF146	EF101	EF101	EF101

Die Variable ist für die Mikrozensen ab 1996 nicht verfügbar, es besteht aber Vergleichbarkeit mit anderen Variablen. Diese sind in der Tabelle rot markiert.


HOME GESIS GML

Dieser Inhalt ist unter einer [Creative Commons-Lizenz](#) lizenziert.



Zentrum für Umfragen,  
Methoden und Analysen

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung

### b) Zugang über Variablenlisten

Wenn sich der Nutzer für einen speziellen Jahrgang des Mikrozensus interessiert und eine Übersicht über die darin enthaltenen Merkmale benötigt, kann er über die Erhebungszeitpunkte eine Erhebung auswählen und sich die Variablenliste ausgeben lassen, die dann zu den Variableninformationen führt. Innerhalb dieser Liste können die Merkmale nach Variablennamen (EF-Nummer), nach der Reihenfolge des Fragebogens und alphabetisch geordnet werden. Die Variablenliste ist somit ein sehr schneller und unkomplizierter Zugang zu variablenbezogenen Informationen:

**Abbildung 6: Zugang über Variablenliste**

**Mikrodaten-Informationssystem**

**MISSY** Mikrozensus: Variablen für MZ1997

HOME OESIS GML

**Variableninformationen**

- Thematische Gliederung
- Variablenliste
- Variablen-Zeitpunkte-Matrix

**Allgemeine Informationen**

- Erhebung
- Daten
- Klassifikationen
- Literatur

**Suche**

Links  
FAQ  
Impressum

Variable	Fragenr.	Variablenlabel
EF1		Bundesland
EF3		Systemfreie Nr.: Auswahlbezirk
EF4		Systemfreie Nr.: Haushalt im Auswahlbezirk
EF5		Laufende Nr.: Person im Haushalt
EF9		Bogenart
EF22		Haushalte in der Wohnung: Anzahl
EF28		Laufende Nr.: Familie im Haushalt
EF30		Alter
EF32	F7	Geschlecht
EF33	F8	Geburtsjahr
EF35	F10	Familienstand
EF36	F10a	Eheschließungsjahr (FB)
EF37	F12	Ehe-, Verwandtschafts- oder Verschwägerungsbeziehung mit Haushaltbezugsp.
EF38	F12a	Verwandtschaftszusammenhang
EF39	F13	Lebenspartner der Haushaltsbezugsp. (FB)
EF40	F13a	Stellung zum Lebenspartner der Haushaltsbezugsp. (FB)
EF41	F14	Weitere Wohnung
EF42	F14a	Hiesige Wohnung: Hauptwohnung
EF43	F15	Deutsche Staatsangehörigkeit
EF44	F15a	1. Staatsangehörigkeit
EF45	F15a	2. Staatsangehörigkeit
EF52		Staatsangehörigk. typisiert
EF53	F17	Zuzugsjahr (FB)
EF70	F18	Kindergarten, -krippe, -hort: gegenwärtiger Besuch
EF71	F19	Schule: gegenwärtiger Besuch
EF72	F19a	Art der besuchten Schule
EF96	F22	Erziehungsurlaub, (Sonder-)Urlaub o.ä. (Berichtsw.)
EF97	F23	Mithelfender Familienangehöriger oder Gelegenheitstätigkeit (Berichtsw.)
EF98	F24	Geringfügige Beschäftigung (Berichtsw.)
EF99	F24a	Geringfüg. Beschäftigung (Berichtsw.): einzige oder hauptsächl. Erwerbstätigk.
EF100	F26	Frühere Erwerbstätigkeit (USP)

### c) Zugang über Variablen-Zeitpunkt-Matrix

Die Variablen-Zeitpunkt-Matrix erfüllt mehrere Funktionen: Zunächst stellt sie einen sehr schnellen Zugang zu variablenbezogenen Informationen auf einer Seite dar und gibt außerdem eine thematisch gegliederte Übersicht über alle in den Scientific Use Files von 1989 bis 2004 erhobenen Merkmale. Ein weiteres Ziel der Matrix besteht darin, auf einen Blick Informationen über die Vergleichbarkeit von Merkmalen bereitzustellen: Variablenbrüche sind in der Matrix rot gekennzeichnet; zudem werden über Infobuttons Informationen über die Art des Variablenbruchs und eine eventuelle Anpassung der Merkmale angeboten. Kann Vergleichbarkeit hergestellt werden, wird zu den entsprechenden SPSS-Routinen verlinkt:

**Abbildung 7: Ausschnitt aus Variablen-Zeitpunkt-Matrix**

Deutscher Lebenspartner der Haushaltsbezugsp.	EF607	EF607	EF607	EF607	EF607	EF607	EF607	EF607	EF607					
<b>*** Familienstand</b>														
Familienstand	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1993	1991	1989	
Familienstand: Haushaltsbezugsp.	EF35	EF35	EF35	EF35	EF35	EF35	EF35	EF35	EF35	EF38	EF38	EF38	EF38	
Familienstand: Familienbezugsp.	EF560	EF560	EF560	EF560	EF560	EF560	EF560	EF560	EF560	EF190	EF190	EF190	EF190	
Familienstand: Lebenspartner der Haushaltsbezugsp.	EF595	EF595	EF595	EF595	EF595	EF595	EF595	EF595	EF595	EF213	EF213	EF213	EF213	
Familienstand: Lebenspartner der Haushaltsbezugsp.	EF658	EF658	EF658	EF658	EF658	EF658	EF658	EF658	EF658					
<b>*** Eheschließungsjahr</b>														
Eheschließungsjahr	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1993	1991	1989	
Eheschließungsjahr: Familienbezugsp.	EF36	EF36	EF36	EF36	EF36	EF36	EF36	EF36	EF36	EF39	EF39	EF39	EF39	
Eheschließungsjahr: Familienbezugsp.	EF596	EF596	EF596	EF596	EF596	EF596	EF596	EF596	EF596	EF214	EF214	EF214	EF214	
<b>*** Staatsangehörigkeit</b>														
Staatsangehörigkeit	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1993	1991	1989	
Staatsangehörigkeit	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF41	EF41	EF41	EF41	
Deutsche Staatsangehörigkeit	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF41	EF41	EF41	EF41	
1. Staatsangehörigkeit	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF41	EF41	EF41	EF41	
2. Staatsangehörigkeit	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF41	EF41	EF41	EF41	
Staatsangehörigkeit: Ha	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF41	EF41	EF41	EF41	
Staatsangehörigkeit: Fa	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF41	EF41	EF41	EF41	
Staatsangehörigkeit: Eh	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF43	EF41	EF41	EF41	EF41	
Staatsangehörigkeit: Lebenspartner der Haushaltsbezugsp.	EF660	EF660	EF660	EF660	EF660	EF660	EF660	EF660	EF660					
<b>*** Bevölkerungstyp</b>														
Bevölkerung: Haupt- oder Nebenwohnsitz	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1993	1991	1989	
Bevölkerung: Haupt- oder Nebenwohnsitz	EF505	EF505	EF505	EF505	EF505	EF505	EF505	EF505	EF505	EF26	EF26	EF26	EF26	
Bevölkerung: Privathaushalte	EF506	EF506	EF506	EF506	EF506	EF506	EF506	EF506	EF506	EF27	EF27	EF27	EF27	
Bevölkerung: Familienwohnsitz	EF508	EF508	EF508	EF508	EF508	EF508	EF508	EF508	EF508	EF28	EF28	EF28	EF28	
Bevölkerung: Wohnsitz der Lebensgem. (Konzept der Lebensformen)	EF513	EF513	EF513	EF513	EF513	EF513	EF513	EF513	EF513					
Bevölkerungsgruppe: Haushaltsbezugsp.	EF569	EF569	EF569	EF569	EF569	EF569	EF569	EF569	EF569	EF199	EF199	EF199		
<b>*** Regionalangaben</b>														
Bundesland	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1993	1991	1989	
Bundesland	EF1	EF1	EF1	EF1	EF1	EF1	EF1	EF1	EF1	EF1	EF1	EF1	EF1	
Gemeindegrößenklasse	EF708	EF708	EF708	EF708	EF708	EF708	EF708	EF708	EF708	EF8	EF8	EF8	EF8	
<b>*** Ausländer und Migration</b>														
Zuzugsjahr	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1993	1991	1989	
Zuzugsjahr	EF53	EF53	EF53	EF53	EF53	EF53	EF53	EF53	EF53	EF47	EF47	EF47	EF47	
Ehegatte im Heimatland	EF54				EF54				EF54	EF48	EF48	EF48	EF48	
Eltern im Heimatland: für Ledige	EF55				EF55				EF55	EF49	EF49	EF49	EF49	
Kinder (unter 6 Jahre): Anzahl im Haushalt	EF56				EF56				EF56	EF50	EF50	EF50	EF50	

Inhalt und Verlinkung der Infobuttons sind mit den Angaben zur Vergleichbarkeit auf Ebene der Variableninformationen identisch. Die Doppelung der Informationen diente hier dem Zweck, wichtige Angaben für das Arbeiten mit den Daten auf verschiedenen Ebenen bereitzustellen. Gerade bei zeitvergleichenden Forschungsansätzen stellt die Variablen-Zeitpunkt-Matrix eine erhebliche Arbeitserleichterung für den Nutzer dar.

#### d) Zugang über Freitextsuche

Eine weitere Zugangsmöglichkeit besteht in der Freitextsuche, die über die google-Suchmaschine realisiert wurde. Dieses Suchfeld ist auf jeder MISSY-Seite zugänglich und ermöglicht eine Stichwortsuche über alle Seiten und Inhalte in MISSY.

Aufgrund der Tatsache, dass MISSY in absehbarer Zeit in das sozialwissenschaftliche Informationsportal SOWIPORT eingebunden werden soll, erschien eine weitere Differenzierung der Suchmöglichkeiten in MISSY nicht mehr notwendig. Innerhalb des SOWIPORTS wird die Suche über das Programm FastSearch realisiert, das z.B. Crosskonkordanzen unterstützt, d.h. dass nicht nur nach dem exakten Suchbegriff, sondern nach inhaltlich äquivalenten Begriffen oder englischen Übersetzungen gesucht werden kann.

### 3.3 Studienbezogene Informationen

Die Studieninformationen sind ebenso wie die Zugänge zu den variablenbezogenen Informationen über die Navigationsleiste auf jeder Seite des Informationssystems direkt abrufbar. Basis für das in MISSY bereitgestellte Angebot bilden die Internetseiten des GML, weshalb im Folgenden nur die Informationen erläutert werden, die für MISSY neu erstellt wurden:

- Für die allgemeine *Studienbeschreibung* wurden die bisher im GML-Web angebotenen Seiten überarbeitet und in DDI 2.0 umgearbeitet. Die neue Version wurde für das GML-Webangebot übernommen.
- Das *Interviewerhandbuch* für den Mikrozensus besteht aus drei Teilen, von denen der zweite Teil, der Informationen für den Interviewer zu den Fragen enthält, für Forscher interessant ist und in MISSY bereitgestellt wird. Dieses Handbuch war den Forschern bislang nicht online zugänglich.
- *Häufigkeitsauszählung*: Es wird eine ungewichtete Häufigkeitsauszählung aller Merkmale als pdf-Dokument zur Verfügung gestellt.
- *Tabellen*: Unter dem Schlagwort „Tabellen“ sind nicht nur die zur Überprüfung des Setups notwendigen Randverteilungen abrufbar, sondern es werden zu den Schwerpunktthemen des Mikrozensus – Erwerbstätigkeit, Bildung und Qualifikation, Haushalts- und Familienformen etc. – systematisch Tabellen und Grafiken zur Verfügung gestellt. Diese Tabellen sollen einerseits einen Eindruck vermitteln, welches Themenspektrum der Mikrozensus abdecken kann, und bieten andererseits relevante wissenschaftliche Informationen zu den verschiedenen Themengebieten. Zusätzlich werden anhand der Tabellen ausgewählte methodische Aspekte vorgestellt und beispielhaft Analysepotenziale des Mikrozensus erläutert. Basis der Tabellen bilden Arbeitsergebnisse des GML, wie schriftliche Publikationen oder Beispiele aus den Workshops. Ziel ist es, ein niedrigschwelliges Angebot zu schaffen, um die Beschäftigung mit den Forschungsmöglichkeiten des Mikrozensus zu forcieren und Anregungen für mögliche Forschungsfragen zu liefern. Diese Tabellen sind auch über die thematische Gliederung abrufbar.
- Unter der Rubrik Verarbeitungshinweise sind neben anderen, zur Datennutzung relevanten Informationen wie z.B. dem Stichprobenfehler *Informationen zum Umgang mit Variablenbrüchen* abgelegt. Die Darstellung inhaltlicher Brüche bei Merkmalen über verschiedene Erhebungszeitpunkte war ein wichtiger Bestandteil des Projekts (vgl. Granato



2001: 39ff). An dieser Stelle können SPSS-Routinen abgerufen werden, die Vergleichbarkeit zwischen Merkmalen herstellen, bei denen es durch die Änderungen des Erhebungsprogramms ab 1996 zu inhaltlichen Brüchen gekommen ist. Zusätzlich kann hier auf eine allgemeine Anleitung zur Herstellung von Vergleichbarkeit zwischen Variablen zugegriffen werden. Die Liste mit SPSS-Routinen ist ebenfalls über die Variablen-Zeitpunkt-Matrix und die Informationsseite der entsprechenden Variablen zugänglich.

- Unter der Rubrik *Konzepte und Definitionen* werden die verschiedenen Konzepte, die dem Mikrozensus zu Grunde liegen, und ihre Veränderungen dokumentiert. Zudem werden die Konzepte zu Haushalt, das traditionelle Familienkonzept und das Konzept der Lebensformen eingehend erläutert. Ebenfalls unter dieser Rubrik abrufbar sind Beschreibungen zur Generierung von *Bandsatzerweiterungen und Typisierungen*. Die Generierung der Variablen wird schriftlich und formal erläutert und wurde im Rahmen des Projektes erarbeitet:

**Abbildung 8: Generierung der Bandsatzerweiterung EF592**

**MISSY** Mikrodaten-Informationssystem

HOME GESIS GML

**Mikrozensus Grundfile MZ1997, Generierung der Variable EF592**

Klassifikationen > Konzepte und Definitionen > Liste generierter Variablen > Typisierungen und Bandsatzerweiterungen im MZ1997 > EF592

**Variableninformationen**

- Thematische Gliederung
- Variablenliste
- Variablen-Zeitpunkte-Matrix

**Allgemeine Informationen**

- Erhebung
- Daten
- Klassifikationen
- Literatur

Suche

Links

FAQ

Impressum

Für jede Familie in Privathaushalten (EF506=1) wird die Familienbezugsperson (EF509=1) ausgewählt. Anschließend wird das Geschlecht der Familienbezugsperson ermittelt (EF32). Es wird eine neue Variable generiert (EF592), die diese Information über das Geschlecht der Familienbezugsperson enthält. Die Zuordnung dieser Information zu allen Mitgliedern einer Familie erfolgt über die Kombination aus Haushaltsnummer (EF1-4) und laufender Nummer der Familie im Haushalt (EF28). Personen, die nicht in Privathaushalten leben, haben bei der generierten Variable (EF592) keinen gültigen Wert.

Die Variable basiert auf dem traditionellen Familienkonzept.

Wenn (EF506=1) & (EF509=1) ⇒ Angabe aus EF32 in EF592 bei allen Personen mit identischer Haushaltsnummer (EF1-4) und identischer Familiennummer (EF28).

Basierend auf: Statistisches Bundesamt (1998): Aufbereitung für den Mikrozensus ab 1996. II Signier- und Plausibilitätskontrolle, Typisierungen und Bandsatzerweiterungen. II.4 Typisierungen und Bandsatzerweiterungen. Unveröffentlichtes Dokument.

Dieser Inhalt ist unter einer [Creative Commons-Lizenz](#) lizenziert.

**ZUMA** Zentrum für Umfragen, Methoden und Analysen

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

- Zusätzlich wurden die Schlagworte des Definitionskatalog und die für die Generierung benötigten Variablen verlinkt, sodass alle zum Verständnis benötigten Informationen auch über die Seiten zur Generierung schnell erreichbar sind.
- Die Literaturliste auf Basis der Mikrozensus Scientific Use Files ist in MISSY – ebenso wie die Tabellen – auch über die thematische Gliederung zugänglich. Hier können für die thematischen Gliederungspunkte auf der ersten Ebene einzeln abgerufen werden, so dass der Nutzer schnell einen Überblick darüber bekommt, welche Veröffentlichungen auf Basis des Mikrozensus es zu dem jeweiligen Themengebiet bereits gibt. Auch beim Zugang über die Studieninformationen ist jetzt auswählbar, ob die Literaturliste nach Themengebieten oder alphabetisch geordnet erscheinen soll.

### 3.4 Usability-Studie

Im Anschluss an die Fertigstellung der Datenbank für die Jahrgänge 1995 und 1997 wurde im zweiten Halbjahr 2006 eine Studie durchgeführt, in deren Rahmen Nutzer sowie externe Experten das System testen und auf seine intersubjektivität überprüfen konnten. Eine Nutzerstudie bot sich als erster Schritt vor der geplanten und beim BMBF beantragten Erweiterung von MISSY um weitere Grundfiles und das neue Regionalfile an, da zu diesem Zeitpunkt noch ohne größeren Aufwand Modifizierungen des Systems vorgenommen werden konnten. Bevor weitere Datensätze in das System integriert werden, ist MISSY somit ein von Nutzern geprüftes und an die Bedürfnisse der Nutzer optimal angepasstes System.

Ziel der Studie war es, Hinweise auf mögliche Probleme und Schwachstellen sowie direkte Verbesserungs- und Gestaltungsvorschläge zu erhalten. Analog zur inhaltlichen Schwerpunktsetzung des Systems war das Hauptanliegen der Studie, die *inhaltlichen Elemente* von MISSY zu testen. Dabei erschien es notwendig, das Informationsangebot sowie die Aufbereitung der Informationen einer kritischen Beurteilung durch die Nutzer zu unterziehen. Über die Evaluation dieser inhaltlichen Aspekte des Informationssystems hinaus wurden außerdem Navigationsstruktur und Informationsdarstellung – und damit die Benutzungsfreundlichkeit des Systems – getestet.

Bei der Nutzerstudie wurde eine Kombination von objektiven und subjektiven Testmethoden angewandt<sup>8</sup>: Zur Gewinnung objektiver Daten bot sich eine kontrollierte Untersuchung an, bei der die Nutzer beim Lösen von Testaufgaben beobachtet wurden. Als Beobachtungsinstrument wurde die Software Morae<sup>9</sup> eingesetzt, mit der das gesamte Geschehen am PC wie z.B. Seitenaufrufe, Mausklicks und -bewegungen sowie Gesichtsausdruck und Audiokommentare aufgezeichnet und analysiert werden konnte. Eine Befragung mittels eines standardisierten Fragebogens gab darüber hinaus Auskunft darüber, wie MISSY von den Nutzern in Hinblick auf Effizienz und Effektivität sowie Informationsgehalt subjektiv beurteilt wird, wobei die Ergebnisse als wichtige Indikatoren für die Zufriedenstellung der Nutzer interpretiert werden können. Dieser standardisierte Fragebogen war in modifizierter Form auch über MISSY onsite verfügbar, so dass auch Besucher der MISSY-Seiten die Möglichkeit hatten, das System zu bewerten.

Als Testpersonen wurden einerseits erfahrene Mikrozensus-Nutzer, andererseits Studierende der Sozial- und Wirtschaftswissenschaften rekrutiert. Insgesamt wurde das System von je zehn Nutzern und zehn Studierenden bewertet. Darüber hinaus wurde MISSY von externen Experten evaluiert<sup>10</sup>.

Eine Reihe von Modifikationen, die als Anregungen aus der Usability-Studie resultierten, konnten in MISSY bereits umgesetzt werden, so z.B.:

---

<sup>8</sup> Einen Überblick über die Methoden zur Evaluation von Software und Informationssystemen bietet Hegner 2003.

<sup>9</sup> <http://www.techsmith.com/morae.asp>

<sup>10</sup> Expertengespräche wurden mit Wendy L. Thomas vom Minnesota Population Centre, Dr. Maximilian Stempfhuber vom IZ, Köln sowie Mitarbeitern der Mikrozensus-Fachabteilung des Statistischen Bundesamtes in Bonn (Gruppe VIII C - Mikrozensus, Arbeitskräftestichprobe, Haushalt und Familie) geführt.

- Verlinkung von Variableninformationen zu den Definitionen
- Verlinkungen bei den Erläuterungen zur Filterführung in den Variableninformationen zu den angegebenen Filtervariablen
- Verlinkungen bei den Angaben zur Generierung von Bandsatzerweiterungen und Typisierungen auf die bei der Generierung relevanten Variablen
- Überarbeitung einiger erläuternder Texte z.B. der Einleitungstext zur Variablen-Zeitpunkt-Matrix
- Übernahme der Informationen zu Variablenbrüchen auf die Ebene der Variableninformationen
- Überarbeitung der Struktur der Seiten zu Variableninformationen
- Einführung einer Blätterstruktur innerhalb der Variableninformationen
- Erleichterung der Navigation durch Pfadangaben auf den Seiten der Studieninformationen und farbige Markierungen zur besseren Unterscheidung von variablen- und studienbezogenen Informationsseiten
- Umsetzung einer größeren Variabilität beim Zugang über Variablenlisten

Aufbau, Ablauf und differenzierte Ergebnisse der Usability-Stude werden im ZUMA-Methodenbericht 2007/01 behandelt.

### 3.5 Technische Umsetzung

Für MISSY wurde der Einsatz von verbreiteten Software-Komponenten angestrebt. Zudem sollte die Software, wenn möglich, aus dem Open-Source-Bereich kommen, da verbreitete Open-Source-Software meist einen hohen Grad an Stabilität erreicht und einer großen User Community zur Verfügung steht. Zudem ist Open-Source-Software für den Einsatz im nicht-kommerziellen Bereich oft kostenlos. Die verschiedenen Ziele bei der Auswahl von Software waren die Stabilität, die Verbreitung, die Portierbarkeit und die Unabhängigkeit von einzelnen Herstellern. Die Portierbarkeit eines Systems ist wichtig, um z.B. ohne große Schwierigkeiten das Betriebssystem oder das Datenbanksystem wechseln zu können. Auf die einzelnen Software-Produkte und ihre Funktionen wird im Folgenden näher eingegangen. Anschließend wird kurz auf Abweichungen der technischen Umsetzung vom Projektantrag hingewiesen.

Bei der Umsetzung wurde eine weitgehende Orientierung an technische Standards angestrebt. Das betrifft im Einzelnen die Datenbanksprache SQL und die Seitenbeschreibungssprachen HTML und CSS:

- Unterschiedliche relationale Datenbanksysteme verwenden oft unterschiedliche SQL-Dialekte. Der umfassende Gebrauch eines SQL-Dialekts hätte eine feste Bindung an ein Datenbanksystem zur Folge. Die eingesetzten SQL-Kommandos orientieren sich am Standard SQL-92. Bei der Erstellung der Datenbankinhalte und der Abfrage davon werden nur SQL-Kommandos eingesetzt, die strikt SQL-92 entsprechen. Die einzige Ausnahme sind längere Bezeichnernamen.
- Die HTML-Seiten entsprechen ganz der aktuellen Version XHTML 1.0. Inkompatible Seiten zu den verschiedenen Web-Browsern werden so weit wie möglich ausge-

geschlossen. Eine einfache technische Weiterverarbeitbarkeit dieser Seiten ist damit auch gewährleistet, weil XHTML dem XML-Standard folgt und somit durch Programme strukturiert gelesen werden kann.

Für die *Eingabe von neuen Informationen* wurden Formulare in InfoPath entwickelt. InfoPath ist Teil der Microsoft Office Suite. Vorteilhaft für die Benutzung bei der Informationseingabe ist die ähnliche Bedienungs Oberfläche wie z.B. Word.

Bei den Formularen können Felder entsprechend konfiguriert werden, so dass z.T. eine Eingabekontrolle stattfinden kann. Außerdem besteht die Möglichkeit der Definition von Textfeldern mit Formatierungsangaben, was z.B. bei der Eingabe von Fragetexten zur Markierung von hervorgehobenen Textstellen relevant war. Die eingegebenen Informationen werden in einer XML-Datei abgelegt, die dann strukturiert weiterverarbeitet werden kann.

Für die *Erstellung des Datenbankschemas* wurden zunächst modellhaft verschiedene graphische Entwicklungswerkzeuge eingesetzt. Ein fertiges Modell wurde dann über den im nächsten Absatz beschriebenen Weg zum Datenbankschema realisiert.

Die Definition eines Datenbankschemas über SQL-Anweisungen, die strikt SQL-92 oder SQL:1999 entsprechen, ist meist nicht möglich, weil gerade dabei die Datenbanksysteme unterschiedliche SQL-Dialekte verwenden. Um dennoch eine vom Datenbanksystem unabhängige Herangehensweise zu gewährleisten, wird das Werkzeug DDLUtils (Data Definition Language Utilities der Apache Group) eingesetzt. Dabei wird das Datenbankschema unabhängig von SQL in einer XML-Datei konfiguriert. Diese Datei wird über ein InfoPath-Formular erzeugt (das Formular ist eine Ergänzung von MISSY zu DDLUtils). Das Werkzeug erzeugt entsprechenden SQL-Dialekt für das gewählte Datenbanksystem.

Für den *Import und die Aktualisierung der Informationen in der Datenbank* wurden die heterogenen und unterschiedlich formatierten Informationsquellen in SQL-Skripte (INSERT / UPDATE) konvertiert. Ein Beispiel dafür ist die Häufigkeitsauszählung aus SPSS. Hier wurde das Ausgabeformat SPSSOMS XML genutzt, um daraus entsprechende SQL-Skripte zu erzeugen.

Die Ausführung der SQL-Skripte erfolgt durch eine zentrale Steuerung. Dafür wird das Java-basierte (damit plattform-unabhängige) Werkzeug Ant der Apache Group eingesetzt. Eine Definition von Aufgaben wurde in einer XML-Konfigurationsdatei pro Verzeichnis vorgenommen. Durch diesen Aufbau können alle Informationen oder nur Teile in die Datenbank eingefügt bzw. darin aktualisiert werden. Ant wird sonst üblicherweise bei der Programmentwicklung eingesetzt.

Das Informationssystem setzt sich aus unterschiedlichen Komponenten zusammen: Die strukturierten Informationen sind in einer relationalen Datenbank gespeichert. Texte und Originaldokumente liegen als HTML- und PDF-Dateien vor. Nutzeranfragen von einem Web-Browser lösen in einem Applikations-Server eine dynamische Zusammenstellung der Informationen aus diesen beiden Informationsquellen aus, wie in Abbildung 9 verdeutlicht wird.

Ein Server unter Windows 2000 wird eingesetzt. Windows wurde gewählt, weil das ursprünglich eingesetzte Datenbanksystem Oracle bei ZUMA nur für Windows zur Ver-

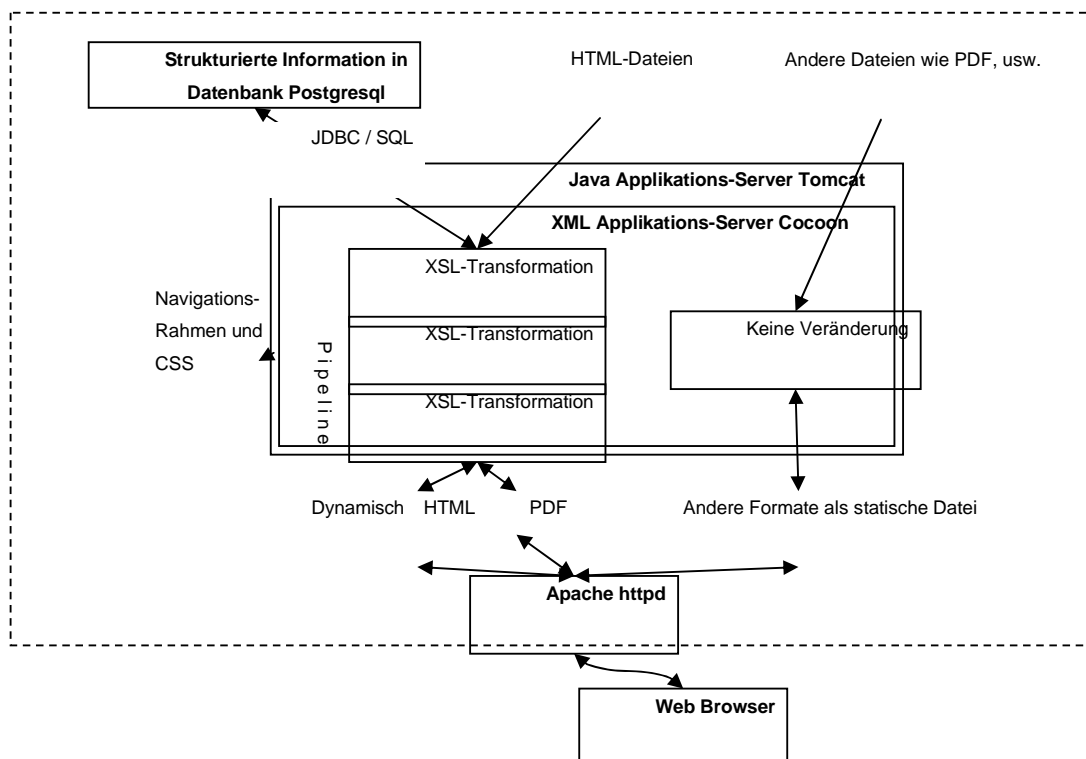
fügung stand. In Zukunft wird Linux als Betriebssystem eingesetzt werden wie sonst auch bei anderen Servern bei ZUMA.

Als *relationales Datenbanksystem* wird das Open-Source-Datenbanksystem Postgresql eingesetzt. Postgresql hat den Vorteil, dass es weitgehend konform zu den ANSI-SQL-Standards ist. Es ist als stabiles Datenbanksystem bekannt und kann auf Linux wie auch Windows eingesetzt werden.

Als *XML-Applikations-Server* wird Cocoon eingesetzt. Es läuft als Java-Servlet im Java-Applikationsserver Tomcat. Als *Web-Server* wird Apache httpd verwendet. Alle drei Komponenten sind Open-Source-Software der Apache Group und stehen für Linux und Windows zur Verfügung.

Cocoon kann Daten aus heterogenen Informationsquellen verarbeiten; hier die Ergebnisse von Datenbankabfragen und Dateien aus dem Dateisystem. Die Ergebnisse der SQL-Datenbankabfragen werden in Cocoon in ein XML-Format umgewandelt; genauso sind die meisten Dateien in XML/XHTML. Durch einen mehrstufigen Prozess (Pipeline) werden diese XML-Informationen dynamisch in eine einheitliche Struktur transformiert. Ein zentral konfigurierbarer Navigationsrahmen wird dabei den HTML-Seiten hinzugefügt. Die Informationen in XML können statt zu HTML-Seiten zum FO-Format (Formatting Objects) transformiert werden; daraus können PDF-Dateien erstellt werden. Dies wird eingesetzt z.B. für große Ausgaben wie das Codebuch.

**Abbildung 9: Technischer Aufbau von MISSY**



Die Datenbankabfragen werden in der Cocoon-eigenen Sprache ESQL formuliert. ESQL ist Teil von XSP (Extended Server Pages). Aufgrund der in ESQL eingebetteten SQL-Kommandos wird dynamisch ein Java-Programm generiert. Dieses Programm richtet die Abfragen an die relationale Datenbank über die Java-Datenbankschnittstelle JDBC.

Apache httpd wird an Stelle von Tomcat als Web Server genutzt, um zusätzliche Funktionen nutzen zu können, wie Kompression von HTML-Seiten zur Einsparung von Bandbreite.

Für die erzeugten HTML-Seiten wird zur Formatierung CSS benutzt. So liegt ein klare Trennung vor zwischen einerseits den Inhalten und dem strukturellen Aufbau in den HTML-Seiten und andererseits dem Seitenlayout, welches in CSS definiert wird. Auf diese Weise können an zentraler Stelle Layoutänderungen vorgenommen werden.

Zusätzlich wird für manche Funktionen auf den HTML-Seiten EcmaScript benutzt, die Browser-unabhängige Version von Javascript.

Im Laufe des Projekts haben sich in technischer Hinsicht folgende Abweichungen vom Projektantrag ergeben:

- Die geplante und auch zuerst eingesetzte Software von Oracle wurde abgelöst durch Open-Source-Komponenten. ZUMA hat während der Projektlaufzeit den Vertrag mit Oracle aus Kostengründen gekündigt. Oracle bietet für Forschungsinstitutionen keine entsprechenden Rabatte. Das Datenbanksystem von Oracle wurde ersetzt durch PostgreSQL. Der Applikations-Server von Oracle wurde ersetzt durch Tomcat und Cocoon der Apache Group.
- Die gewählte Lösung setzt sich zusammen aus stabilen und verbreiteten Open-Source-Komponenten. Tomcat und Cocoon werden bei ZUMA auch bei anderen Anwendungen wie dem Wohlfahrtssurvey-Online eingesetzt. Der Einrichtungs- und Verwaltungsaufwand dieser Komponenten ist im Gegensatz zu Oracle gesunken. Zudem ist die eingesetzte Open-Source-Software kostenlos.
- Ein ursprünglich optional geplantes Java-Applet als zusätzliche Benutzerschnittstelle wurde nicht realisiert. Die Web-Browser haben sich in den letzten Jahren zu einer stabilen und vielfältigen Plattform für Benutzungsoberflächen entwickelt. Die Anforderungen von MISSY konnten gut in HTML-Seiten in Kombination mit EcmaScript realisiert werden. Zudem hat der Einsatz von Java-Applets keine große Verbreitung erfahren. Aus diesen Gründen wurde von einer derartigen Entwicklung abgesehen.

### 3.6 Nutzen und Verwertbarkeit

Das langfristige Ziel des MISSY-Projekts besteht im Aufbau eines umfassenden Informationssystem für den Mikrozensus, das langfristig die Service-Seiten des German Microdata Labs ablösen soll. Als nächster Schritt ist die Einbindung von MISSY in den SOWIPOINT geplant. SOWIPOINT soll als Fachportal für die Sozialwissenschaften alle Produkte der GESIS und weiteren sozialwissenschaftlichen Instituten beinhalten und somit als themenübergreifendes Recherche- und Informationsportal für die sozialwissenschaftliche Profession etabliert werden. Zur Erweiterung von MISSY um die Metadaten weiterer Scientific Use Files des Mikrozensus und für eine Einbindung von Informationen zum demnächst bereitgestellten

Regionalfile des Mikrozensus wurde im Dezember 2006 ein Folgeantrag zur Finanzierung beim BMBF gestellt.

### 3.7 Fortschritt bei anderen Stellen

Neben dem Mikrodateninformationssystem MISSY gibt es weitere Informationssysteme, deren Angebot ebenfalls Metadaten zum Mikrozensus umfasst. Online verfügbar sind derzeit drei Systeme: Das Informationssystem GENESIS (Gemeinsames neues statistisches Informationssystem) der Statistischen Ämter des Bundes und der Länder, das „FDZ-Metadatensystem Online“ der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder und das Metadatensystem für arbeitsmarktforschungsrelevante Datensätze des Forschungsinstitut zur Zukunft der Arbeit (IZA). Diese Informationssysteme weisen eine wichtige Gemeinsamkeit auf: Sie dienen vor allem dazu, sich einen ersten Überblick über verschiedene Datenquellen zu verschaffen. Zu diesem Zweck bieten sie allgemeine Informationen über den Mikrozensus und darüber hinaus zum Teil auch die Merkmalsausprägungen einzelner Variablen. MISSY können sie allerdings nicht ersetzen, da MISSY stärker analyseorientiert ist und sowohl umfassende Informationen als auch detaillierte Angaben z.B. auf Ebene der Variablen bereitstellt: Hinweise für das Arbeiten mit den Daten, Literaturangaben zu Forschungsergebnissen mit dem Mikrozensus, Übersichten über alle erhobenen Merkmale (Variablen-Zeitpunkte-Matrix) mit Angaben über Variablenbrüche, Verknüpfung zwischen verwandten Themenbereichen etc. Da die anderen Informationssysteme inhaltlich breit angelegt sind und verschiedene Datensätze dokumentieren, kann die Dokumentation eines einzelnen Datensatzes kaum erschöpfend geleistet werden. MISSY ist zwar generell offen für die Einbindung anderer Datensätze, konzentriert sich aber in den bisherigen Phasen auf Metadaten zum Mikrozensus und verfolgt damit ein anderes Ziel: das der möglichst vollständigen Dokumentation, die sich nicht auf die unmittelbaren und grundlegenden Metadaten beschränkt, sondern alle für die Analyse der Daten relevanten Informationen umfasst, die nach inhaltlichen Gesichtspunkten verknüpft werden. Die sozial- und wirtschaftswissenschaftliche Ausrichtung der Informationsaufbereitung und -verknüpfung ist eine Besonderheit von MISSY, die die anderen Informationssysteme in dieser Form nicht leisten können. Daher stellt MISSY eine unverzichtbare Ergänzung zu den anderen Informationssystemen dar.

### 3.8 Erfolgte/geplante Veröffentlichung des Ergebnisses

Im Gegensatz zu üblichen Forschungsprojekten, deren Resultat aus einem Abschlussbericht besteht, welcher die Grundlage für weitere Veröffentlichungen bilden kann, besteht das Endprodukt dieses Projekts aus einem frei über das Internet verfügbaren System. Das Ergeb-

nis des Projekts kann somit als veröffentlicht gelten. Neben dem über das Web verfügbaren System gibt es weitere Veröffentlichungen:

- Veröffentlichungen in Zeitschriften und Arbeitsberichte

Bohr, Jeanette, 2007: Abschlussbericht MISSY Nutzerstudie. ZUMA-Methodenbericht 2007/01

Bohr, Jeanette et al., 2006: Problems of Comparability in the German Microcensus over Time and the new Version of DDI 3.0, Iassist Quarterly vol 30, 13-19<sup>11</sup>

Janßen, Andrea und Jeanette Bohr, 2007: Microdata Information System MISSY, erscheint in: Iassist Quarterly vol 30-2

- Vorträge im Rahmen des Projektes

Bohr, Jeanette, Andrea Janßen und Joachim Wackerow: Problems of comparability in the German Microcensus over time and the new DDI version 3.0. IASSIST (International Association for Social Science Information Services & Technology) 2006. ICPSR, University of Michigan, Ann Arbor, 23. – 26.5.2006<sup>12</sup>

Janßen, Andrea, Jeanette Bohr und Wackerow, Joachim: Microdata Information System MISSY. IASSIST (International Association for Social Science Information Services & Technology) 2006. ICPSR, University of Michigan, Ann Arbor, 23. – 26.05.2006<sup>13</sup>

Janßen, Andrea und Jeanette Bohr: Datendokumentation amtlicher Mikrodaten – das Mikrodateninformationssystem MISSY. Statistische Woche Dresden, 18. – 21.09.2006

Jeanette Bohr und Andrea Janßen: Serviceangebot des GML: Mikrodaten-Informationssystem MISSY. ZUMA-Workshop „Einführung in die Besonderheiten und neuen Analyse-möglichkeiten des Mikrozensus ab 2005“, Mannheim, 26./27. 04.2007

- Vorstellung des Projektes

Jeanette Bohr und Andrea Janßen: MISSY-Poster im Rahmen der Postersession auf der Konferenz des Rates für Sozial- und Wirtschaftsdaten, Wiesbaden, am 11./12. 05. 2006

Jeanette Bohr: Poster der Usability-Studie im Rahmen der Postersession der GOR-Konferenz (General Online Research). Leipzig, 26-28.03. 2007 in Leipzig

Vorstellung von MISSY in den ZUMA-Nachrichten Nr. 59, 2006

Vorstellung von MISSY im GESIS-Resport 2/07

---

<sup>11</sup> Siehe: <http://iassistdata.org/publications/iq/iqvol30.html>

<sup>12</sup> Siehe: [http://www.iassistdata.org/conferences/2006/presentations/E1\\_GesZUMA.pdf](http://www.iassistdata.org/conferences/2006/presentations/E1_GesZUMA.pdf)

<sup>13</sup> Siehe: [http://www.iassistdata.org/conferences/2006/presentations/C3\\_MISSY.ppt](http://www.iassistdata.org/conferences/2006/presentations/C3_MISSY.ppt)



### **III    Anlage: Veröffentlichungen, Vorträge und Veranstaltungen**

#### **Monographien und Sammelbände**

Dorbritz, Jürgen, Andrea Lengerer und Kerstin Ruckdeschel, 2005: Einstellungen zu demographischen Trends und zu bevölkerungsrelevanten Politiken. Ergebnisse der Population Policy Acceptance Study in Deutschland. Wiesbaden: Bundesinstitut für Bevölkerungsforschung.

Janßen, Andrea und Ayca Polat, 2005: Zwischen Integration und Ausgrenzung. Lebensverhältnisse türkischer Migranten der zweiten Generation. Oldenburg: Universität.

Gestring, Norbert, Andrea Janßen und Ayça Polat, 2006: Prozesse der Integration und Ausgrenzung. Türkische Migranten der zweiten Generation. Wiesbaden: VS Verlag für Sozialwissenschaften.

#### **Veröffentlichungen in Zeitschriften**

Bohr Jeanette, Andrea Janßen und Joachim Wackerow, 2006: Problems of Comparability in the German Microcensus over Time and the New DDI Version 3.0. IASSIST Quarterly vol. 30, S. 13-19.

Jäckel, Michael und Jeanette Bohr, 2004: Jugendkulturelle Stile und Moden. Eine empirische Untersuchung des Modeverhaltens. Deutsche Jugend 9 (52): 388-396.

Janßen, Andrea und Jeanette Bohr, 2007: Microdata Information System MISSY, erscheint in: Iassist Quarterly vol 30-2

Janßen, Andrea und Ayca Polat, 2006: Soziale Netzwerke türkischer Migrantinnen und Migranten. In: Aus Politik und Zeitgeschichte. S. 11-17, Heft 1-2

Lengerer, Andrea, 2004: Familienpolitische Regimetypen in Europa und ihre Bedeutung für den Wandel der Familie. Ein Überblick über den Stand der Forschung. In: Zeitschrift für Bevölkerungswissenschaft, 29, 99-121.

Lengerer, Andrea, 2005: Zur Akzeptanz von Familienpolitik. Zeitschrift für Bevölkerungswissenschaft 29, S. 387-420.

Lengerer, Andrea, Andrea Janßen und Jeanette Bohr, 2007: Familiensoziologische Analysepotentiale des Mikrozensus. Erscheint in: Zeitschrift für Familienforschung, 19, Heft 2

Lengerer, Andrea; Thomas Klein, 2007: Der langfristige Wandel partnerschaftlicher Lebensformen im Spiegel des Mikrozensus. In: Wirtschaft und Statistik, 4/2007, 433-447

Lüttinger, Paul, Holger Breiholz und Yvonne Lechert, 2003: Ergebnisse der zweiten Befragung von Nutzern der Mikrozensus-Scientific Use Files. ZUMA-Nachrichten 53: 128-135. Ebenso erschienen in: Statistisches Bundesamt (Hg.), 2003: Wirtschaft und Statistik. Heft 11/2003: 1055-1058.

Lüttinger, Paul, Bernhard Schimpl-Neimanns, Heike Wirth und Georgios Papastefanou, 2004: The German Microdata Lab at ZUMA: Services Provided to the Scientific Community, S. 455-467 in: Schmollers Jahrbuch („Zeitschrift für Wirtschafts- und Sozialwissenschaften“), 3/2004, Berlin: Duncker & Humblot.

Lüttinger, Paul, Bernhard Schimpl-Neimanns, Heike Wirth und Georgios Papastefanou, 2003: Mikrodaten (German Microdata Lab): Das Servicezentrum für amtliche Mikrodaten bei ZUMA. ZUMA-Nachrichten 52:153-172.

Lüttinger, Paul und Heike Wirth, 2004: Zur Weitergabe von älteren amtlichen Mikrodaten an die Wissenschaft. Allgemeines Statistisches Archiv 88 (4): 473-486.

Schroedter, Julia 2006: Binationale Ehen in Deutschland. In: Wirtschaft und Statistik. Statistisches Bundesamt: Wiesbaden, S. 419-431, 4/2006.

### **Beiträge zu Sammelbänden**

Gestring, Norbert, Andrea Janßen, Ayca Polat und Walter Siebel, 2006: Prozesse der Integration und Ausgrenzung bei türkischen Migranten der zweiten Generation. In: Hartwig Heine, Michael Schumann und Volker Wittke: Wer den Ast absägt, auf dem er sitzt, kann deshalb noch längst nicht fliegen. Innovationen zwischen institutionellem Wandel und Pfadkontinuitäten. 2006, Edition Sigma: Berlin, S. 215-237

Polat, Ayca und Andrea Janßen, 2006: „Jeder Topf findet seinen Deckel und Deutschland ist für uns der Deckel geworden“ – Die zweite Generation türkischer Migranten in Hannover. In: Wolfgang Puschmann und Herbert Schmalstieg: Stadt und Kirche im demografischen Wandel. 2006, Ev.-luth. Stadtkirchenverband: Hannover, S. 129-148.

Shahla, Hossein, 2005: Political attitudes, social structure and voting behaviour in Macedonia. S. 119-134 in: Svetomir Škarić: Founding Elections in Eastern Europe, Democratic Elections in Macedonia 1990-2002. Berlin: Sigma.

### **Arbeitspapiere**

Bohr, Jeanette, 2004: Dokumentation MISSY-Web. ZUMA-Interner Projektbericht 2004/06.

Hubert, Tobias und Helga Christians, 2004: Mikrozensus 1999. Codebuch und Grundauszählung. ZUMA-Interner Projektbericht 2004/17.

Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1973 Codebuch und Grundauszählung: ZUMA-Interner Projektbericht 04/2006.

Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1976 Codebuch und Grundauszählung: ZUMA-Interner Projektbericht 05/2006.

Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1982 Codebuch und Grundauszählung: ZUMA-Interner Projektbericht 06/2006.

Lengerer, Andrea, Jeanette Bohr und Andrea Janßen, 2005: Haushalte, Familien und Lebensformen im Mikrozensus – Konzepte und Typisierungen. ZUMA-Arbeitsbericht 2005/05.

Lengerer, Andrea und Julia H. Schroedter, 2005: Ergebnisse und Empfehlungen aus der Teilnehmerbefragung des ZUMA-Workshops „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“ am 23./24. Februar 2005. German Microdata Lab Arbeitspapier 2005/02.

- Lengerer, Andrea, Julia H. Schroedter, Hossein Shahla und Andreas Fischer, 2005: Mikrozensus Scientific Use File 2003. Codebuch und Grundauszählung. ZUMA-Interner Projektbericht 06/2005.
- Lengerer, Andrea, Julia H. Schroedter, Hossein Shahla und Andreas Fischer, 2005: Mikrozensus Scientific Use File 2002. Codebuch und Grundauszählung. ZUMA-Interner Projektbericht 03/2005.
- Lüttinger, Paul und Yvonne Lechert, 2003: Ergebnisse der zweiten Befragung von Nutzern der Mikrozensus-Scientific Use Files. ZUMA-Interner Projektbericht 2003/17.
- Lüttinger, Paul und Heike Wirth, 2004: Regelungen zur Weitergabe von amtlichen Mikrodaten an die Wissenschaft in den Bundesstatistikgesetzen seit 1953. Arbeitspapier German Microdata Lab, Arbeitspapier 2004/01.
- Lüttinger, Paul und Heike Wirth, 2004: Auswahl relevanter Datenzeitpunkte für die Bereitstellung von Scientific Use Files von Mikrozensen vor 1989. German Microdata Lab Arbeitspapier 2004/02.
- Schroedter, Julia H. und Andrea Lengerer, 2005: Ergebnisse und Empfehlungen aus der Teilnehmerbefragung des ZUMA-Workshops „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“ im September 2004. German Microdata Lab Arbeitspapier 2005/01.
- Shahla, Hossein, Julia Schroedter und Andreas Fischer, 2004: Mikrozensus 2001. Codebuch und Grundauszählung. ZUMA-Projektbericht 22/2004.

### **Methodenberichte**

- Hubert, Tobias, 2005: Vergleich der Mikrozensus Scientific Use Files 1999 und 2000. ZUMA-Methodenbericht 2005/04.
- Kollinger, Nicole und Hossein Shahla, 2004: Vergleich der Mikrozensen 2000 und 2001. ZUMA-Methodenbericht 2004/09.
- Lechert, Yvonne, Julia H Schroedter und Paul Lüttinger, 2006: Die Umsetzung der Bildungsklassifikation CASMIN für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensen 1976-2004. ZUMA-Methodenbericht 2006/12.
- Lechert, Yvonne und Paul Lüttinger, 2006: Volks- und Berufszählung (10 % Stichprobe) Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/11.
- Lechert, Yvonne, Paul Lüttinger und Schunck, Reinhard, 2006: Vergleich der Mikrozensus Scientific Use Files 1976 und 1982. ZUMA-Methodenbericht 2006/10.
- Lechert, Yvonne, Paul Lüttinger und Schunck, Reinhard, 2006: Vergleich der Mikrozensus Scientific Use Files 1973 und 1976. ZUMA-Methodenbericht 2006/09.
- Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1982. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/03.
- Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1976. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/02.

- Lechert, Yvonne und Paul Lüttinger, 2006: Mikrozensus Scientific Use File 1973. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/01.
- Lengerer, Andrea und Mara Boehle, 2006: Rekonstruktion von Bandsatzerweiterungen zu Haushalt, Familie und Lebensformen im Mikrozensus. ZUMA-Methodenbericht 2006/05.
- Lengerer, Andrea, Tobias Hubert und Nicole Kollinger, 2005: Vergleich der Mikrozensus Scientific Use Files 2001 und 2002. ZUMA-Methodenbericht 2005/02.
- Lengerer, Andrea und Hossein Shahla, 2006: Mikrozensus Scientific Use File 2004 - Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2006/04.
- Schroedter, Julia H. und Tobias Hubert, 2005: Vergleich der Mikrozensus Scientific Use Files 2002 und 2003. ZUMA-Methodenbericht 2005/07.
- Schroedter, Julia H., Yvonne Lechert und Paul Lüttinger, 2006: Die Umsetzung der Bildungsskala ISCED-1997 für die Volkszählung 1970, die Mikrozensus- Zusatzserhebung 1971 und die Mikrozensen 1976-2004. ZUMA-Methodenbericht 2006/08.
- Schroedter, Julia H. und Kathrin Leim, 2006: Mikrozensus Scientific Use File 2004 - Vergleich mit dem vorhergehenden Erhebungszeitpunkt. ZUMA-Methodenbericht 2006/07.
- Shahla, Hossein, 2005: Mikrozensus Scientific Use File 2002. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2005/01.
- Shahla, Hossein, 2004: Mikrozensus Scientific Use File 2001. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2004/08.
- Shahla, Hossein; Andreas Fischer und Tobias Hubert, 2005: Mikrozensus Scientific Use File 2003. Dokumentation und Datenaufbereitung. ZUMA-Methodenbericht 2005/06.

### **Mikrodatentools**

- John, Kristina 2006: Rekonstruktion von Bandsatzerweiterungen mittels zeilen- bzw. satz- übergreifender Operationen in STATA.
- Lechert, Yvonne, Julia Schroedter und Paul Lüttinger, 2006: Die Umsetzung der CASMIN-Bildungsskala für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensen 1976-2003. ZUMA-Methodenbericht 2006/12.
- Lengerer, Andrea und Mara Boehle, 2006: Rekonstruktion von Bandsatzerweiterungen zu Haushalt, Familie und Lebensformen im Mikrozensus. ZUMA-Methodenbericht 2006/05).
- Schroedter, Julia H., Yvonne Lechert und Paul Lüttinger, 2006: Die Umsetzung der Bildungsskala ISCED-1997 für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensen 1976-2003. ZUMA-Methodenbericht 2006/08.

## **Workshops/Veranstaltungen**

### **2006**

Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten, Mannheim, 08. – 09.03.2006.

Multivariate Analyse kategorialer Daten. ZUMA, Mannheim, 15. – 16.03.2006.

Ausländische Bevölkerung im Mikrozensus. ZUMA, Mannheim, 05.10.2006.

### **2005**

Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten. ZUMA, Mannheim, 23. – 24.02.2005.

Multivariate Analyse kategorialer Daten. ZUMA, Mannheim, 20. – 21.04.2005.

Internationaler Vergleich von Hintergrundvariablen. ZUMA, Mannheim, 01. – 02.06.2005.

Multivariate Analyse kategorialer Daten. ZUMA, Mannheim, 05. – 06.10.2005.

4. Nutzerkonferenz: „Forschung mit dem Mikrozensus: Analysen zur Sozialstruktur und zum Arbeitsmarkt“, Mannheim, 12. – 13.10.2005.

### **2004**

Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Eine Einführung. ZUMA, Mannheim, 17.02.2004.

Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten. ZUMA, Mannheim, 29. – 30.10.2004.

Einführung in das Arbeiten mit dem Mikrozensus am Beispiel der Bildung und der Erwerbstätigkeit. Einführung in den Mikrozensus. Bundesinstitut für Berufsbildung (BIBB), Bonn, 25.11.2004.

### **2003**

3. Mikrozensus Nutzerkonferenz: Forschung mit dem Mikrozensus: Analysen zur Sozialstruktur und zum Arbeitsmarkt. Mannheim, 09. – 10.10.2003.

## Vorträge

### 2006

- Bohr, Jeanette und Andrea Janßen: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel Arbeitsmarkt- und Wohnsituation: Deutsche und Migranten im Vergleich. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“. Mannheim, 08. – 09.03.2006
- Bohr, Jeanette, Andrea Janßen und Joachim Wackerow: Problems of comparability in the German Microcensus over time and the new DDI version 3.0. IASSIST (International Association for Social Science Information Services & Technology) 2006. ICPSR, University of Michigan, Ann Arbor, 23. – 26.5.2006
- Janßen, Andrea und Jeanette Bohr: Datendokumentation amtlicher Mikrodaten – das Mikrodateninformationssystem MISSY. Statistische Woche Dresden, 18. – 21.09.2006.
- Janßen, Andrea, Jeanette Bohr und Wackerow, Joachim: Microdata Information System MISSY. IASSIST (International Association for Social Science Information Services & Technology) 2006. ICPSR, University of Michigan, Ann Arbor, 23. – 26.05.2006
- Janßen, Andrea und Norbert Gestring: Integration und ethnische Communities - Die Rolle des Quartiers. Treffen des Arbeitskreises „Stadt und Migration“. Centre Marc Bloch, Berlin, 02.12.2006
- Janßen, Andrea und Julia Schroedter: Übungen zu den Auswertungen im Mikrozensus I + II. ZUMA-Workshop „Ausländische Bevölkerung im Mikrozensus“, Mannheim, 05.10.2006
- Lechert, Yvonne: Vergleichbarkeit von Variablen des Mikrozensus über die Zeit. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“, Mannheim, 08. – 09.03.2006.
- Lengerer, Andrea und Julia H. Schroedter: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel Bildung und Einkommen“. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“, Mannheim, 08. – 09.03.2006.
- Lüttinger, Paul: Vergleichbarkeit von Variablen des Mikrozensus über die Zeit. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“, Mannheim. 08. – 09.03.2006.
- Lüttinger, Paul, und Lechert, Yvonne, 2006: Vergleichbarkeit von Variablen des Mikrozensus über die Zeit. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung“. 08. – 09.03. 2006.
- Schroedter, Julia und Andrea Janßen: Ausländische Bevölkerung im Mikrozensus: Möglichkeiten und Grenzen von Auswertungen. ZUMA-Workshop „Ausländische Bevölkerung im Mikrozensus“, Mannheim, 05.10.2006.

Shahla, Hossein: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel Erwerbstätigkeit. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“, Mannheim, 08. – 09.03.2006.

## 2005

Bohr, Jeanette und Andrea Janßen: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel Arbeitsmarkt- und Wohnsituation: Deutsche und Migranten im Vergleich. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“, Mannheim, 23. – 24.02.2005.

Janßen, Andrea: Lebenssituation türkischer Migranten der zweiten Generation. Treffen der Arbeitsgemeinschaft „Stadt und Migration“, Göttingen, 29.10.2005.

Lengerer, Andrea: Podiumsdiskussion „Mit Soziologie in die Praxis“ am Institut für Soziologie der Universität Heidelberg, 20.05.2005

Lengerer, Andrea und Julia H. Schroedter: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel Bildung und Einkommen. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“, Mannheim, 23. – 24.02.2005.

Lüttinger, Paul: Der Mikrozensus: Eine Einführung. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung“. Mannheim, 23.02.2005.

Shahla, Hossein,: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel Erwerbstätigkeit. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Einführung in das Arbeiten mit den Daten“, Mannheim, 23. – 24.02.2005.

## 2004

Bohr, Jeanette: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel Erwerbstätigkeit. ZUMA-Workshop „Einführung in das Arbeiten mit dem Mikrozensus am Beispiel der Bildung und der Erwerbstätigkeit. Einführung in den Mikrozensus“, Bundesinstitut für Berufsbildung (BIBB), Bonn, 25.11.2004.

Bohr, Jeanette und Helga Christians: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel Erwerbstätigkeit und Einkommen. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Arbeiten mit den Daten“, Mannheim, 29. – 30.09.2004

Janßen, Andrea und Ayca Polat: „Es ist ´ne soziale Gegend“ –Benachteiligende Effekte in Migrantenvierteln? 32. Kongress der DGS, München, 5.10.2004

Lengerer, Andrea, 2004: Aspekte des Wandels der Familie und die Rolle der Familienpolitik. Seminar „Kinderwagen – Kinder wagen? Familie und Kinderentwicklung heute und in der Zukunft“ des Arbeitskreises Demokratie der Friedrich-Naumann-Stiftung, Gumpersbach, 24. – 26.09.2004.

Lüttinger, Paul: „Das German Microdata Lab“, Statistische Woche Frankfurt, 22.09.2004.

Lüttinger, Paul, und Bernhard Schimpl-Neimanns: Der Mikrozensus: Eine Einführung. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung“. Mannheim, 17.02.2004.

Schroedter, Julia H. und Hossein Shahla: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel Bildung und Erwerbstätigkeit. ZUMA-Workshop „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Arbeiten mit den Daten“, Mannheim, 29. – 30.09.2004.

Schroedter, Julia H. und Hossein Shahla: Einführung in das Arbeiten mit dem Mikrozensus am Beispiel der Bildung. ZUMA-Workshop „Einführung in das Arbeiten mit dem Mikrozensus am Beispiel der Bildung und der Erwerbstätigkeit. Einführung in den Mikrozensus“, Bundesinstitut für Berufsbildung (BIBB), Bonn, 25.11.2004.

## **2003**

Lüttinger, Paul: „Ergebnisse der zweiten Befragung von Nutzern der Mikrozensus Scientific Use Files. 3. Nutzerkonferenz „Forschung mit dem Mikrozensus“, Mannheim, 09. – 10. 10. 2003.

## **Teilnahme an wissenschaftlichen Tagungen**

### **2006**

3. Konferenz für Wirtschafts- und Sozialdaten. 11.05.2006, Wiesbaden. (Jeanette Bohr, Yvonne Lechert, Andrea Lengerer, Andrea Janßen, Paul Lüttinger, Hossein Shala).

International Association for Social Sciences Information Service & Technology (IASSIST) Conference, Ann Arbor, Michigan, 24.05. 2006. (Jeanette Bohr, Andrea Janßen, Joachim Wackerow).

Statistische Woche, Dresden, 18. – 21.09.2006 (Jeanette Bohr, Yvonne Lechert, Andrea Janßen)

Unternehmen und Arbeitsmarkt in Bewegung – Was gewinnt die Forschung durch amtliche Mikrodaten? Veranstaltet vom Forschungsdatenzentrum, Berlin, 31.08. – 01.09.2006 (Jeanette Bohr, Andrea Janßen)

Kongress der Deutschen Gesellschaft für Soziologie, Universität Kassel, 09. – 13.10.2006 (Jeanette Bohr, Andrea Lengerer, Andrea Janßen, Julia Schroedter)

### **2005**

Amtliche Mikrodaten für die wissenschaftliche Forschung – Regionale Nutzerkonferenz: Bevölkerungs- und Sozialstatistiken, Köln, 06.04.2005 (Andrea Lengerer)

Amtliche Mikrodaten für die wissenschaftliche Forschung – Regionale Nutzerkonferenz: Bevölkerungs- und Sozialstatistiken, Berlin, 21. – 22.04.2005 (Julia Schroedter)

Herbsttagung der DGS-Sektion Stadt- und Regionalsoziologie, Köln 14. – 15.10.2005 (Andrea Janßen)



**2004**

2. Konferenz für Wirtschafts- und Sozialdaten. 04. – 05.06.2004, Wiesbaden. (Paul Lüttinger)

Statistische Woche, Frankfurt, 20. – 23.09.2004 (Paul Lüttinger)

**Auszeichnungen**

Die Diplomarbeit von *Julia H. Schroedter* „Binationale Ehen in Deutschland. Der Einfluss der strukturellen auf die soziale Integration“, die an der Universität Mannheim bei Prof. Dr. Frank Kalter entstand, wurde gemeinsam mit der Diplomarbeit von Stephan Bechtel „Verdienststrukturen in Baden-Württemberg: Eine empirische Analyse mit Hilfe von Quantilsregressionen“ (Universität Hohenheim) im Jahr 2005 mit dem Gerhard-Fürst-Preis ausgezeichnet.

#### **IV Anlage: Erfolgskontrollbericht zum Pilotprojekt zum Aufbau eines Servicezentrums für Mikrodaten**

##### **Beitrag des Ergebnisses zu förderpolitischen Zielen**

Das Pilotprojekt ist eingeordnet in die förderpolitischen Ziele zur Verbesserung des Datenzugangs und der informationellen Infrastruktur für Daten der amtlichen Statistik und folgt somit den Empfehlungen der Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI), insbesondere auch die Nutzbarkeit des Mikrozensus für die Wissenschaft weiterzuentwickeln. Dieses Ziel sollte in drei Arbeitsschwerpunkten erreicht werden: Im ersten Teilbereich war eine Stärkung der Service-, Beratungs- und Forschungsleistungen zum Mikrozensus vorgesehen. Im zweiten Bereich sollte eine Lösung des Problems angestrebt werden, dass die Daten von Mikrozensus, die vor 1989 erhoben worden waren, bis zum Projektbeginn nur in absolut anonymer Form (Aggregatdaten) für die Wissenschaft zur Verfügung standen, und mit diesen Daten elaborierte Analyseverfahren somit nicht durchgeführt werden konnten. Im dritten Teilbereich sollte für die Nutzer von Scientific Use Files des Mikrozensus ein datenbankgestütztes Informationssystem entwickelt werden, mit dem Ziel, die bislang zu jedem Scientific Use File zahlreich vorhandenen und über das Web verfügbaren Metainformationen zusammenzuführen. Die drei sich ergänzenden Teilprojekte tragen in ihrer Summe über die systematische Bereitstellung wissenschaftlicher Serviceangebote zu amtlichen Mikrodaten zu einer erheblich verbesserten Nutzungsgrundlage dieser Daten für die gesamte Wissenschaftsgemeinschaft bei. Sie ermöglichen Forschern den schnellen und effizienten Analysezugang zu den Daten und bewirken erhebliche Einsparungen von Zeit und monetären Ressourcen, da die verschiedenen Grundlagenarbeiten nicht von den Einzelforschern selbst erledigt werden müssen.

##### **Wissenschaftlich-technische Ergebnisse, erreichte Nebenergebnisse und gesammelte wesentliche Erfahrungen**

Die Ergebnisse der Teilprojekte sind ausführlich in den Kapiteln 1.1, 1.2 und 1.3 beschrieben. Im ersten Teilprojekt konnten die Dokumentationsarbeiten für die aktuellen Mikrozensus bis zur Erhebung 2004 zu Ende gebracht werden; im zweiten Teilprojekt gelang es, das für die Weitergabe älterer Mikrozensus grundlegende Rechtsproblem zu lösen und drei Mikrozensus vor 1989 (1973, 1976, 1982) als Scientific Use File für die Wissenschaft verfügbar zu machen; im dritten Teilprojekt wurde das Mikrodaten Informationssystem fertig gestellt und

online für die Nutzer verfügbar gemacht. Zum Zeitpunkt des Projektendes sind die kompletten Informationen zu zwei Mikrozensus in MISSY verfügbar.

### **Fortschreibung des Verwertungsplanes**

Das Projekt hat mit seinen Teilprojekten eng an die bisher im Bereich amtlicher Mikrodaten geleisteten Arbeiten angeknüpft und diese weiter konsolidiert. Eine Fortschreibung des Verwertungsplanes ist vor allem in zwei Bereichen gegeben, zum einen in der weiteren Erschließung älterer Mikrozensus-Erhebungen als Scientific Use File, was nun rechtlich problemlos möglich ist, sowie in der weiteren Konsolidierung des Mikrodateninformationssystems durch das Einbinden weiterer Mikrozensus-Erhebungen in die Datenbank.

### **Arbeiten, die zu keiner Lösung geführt haben**

Die den Projekten zugrunde liegenden Aufgaben und Probleme wurden mit Erfolg gelöst. Einzig im Teilprojekt der Rückerschließung älterer Mikrozensus konnte die Zahl der geplanten Scientific Use Files nicht realisiert werden: Statt der geplanten sieben Files wurden bis zum Projektende nur drei Files für die Wissenschaft zur Verfügung gestellt.

### **Einhaltung der Ausgaben- und Zeitplanung**

Das Pilotprojekt zum Aufbau eines Servicezentrums für Mikrodaten der GESIS beim ZUMA wurde für den Zeitraum vom 1.7.2003 bis 31.6.2006 gefördert und aufgrund der Verzögerungen im Bereich der Rückerschließung älterer Mikrozensus kostenneutral bis zum 31.12.2006 verlängert. Der Projektablauf entsprach, bis auf die zeitlichen Verzögerungen im Bereich Rückerschließung, im Wesentlichen der Planung.

## V Kurzfassung: Berichtsblatt

1. ISBN oder ISSN	2. Berichtsart Abschlussbericht	
3a. Titel des Berichts: Abschlussbericht zum Vorhaben: Verbesserung des Zugangs der Wissenschaft zu Mikrodaten Teilprojekt: Pilotprojekt zum Aufbau eines Servicezentrums für Mikrodaten der GESIS beim ZUMA		
3b. Titel der Publikation		
4a. Autoren des Berichts Andrea Janßen, Jeanette Bohr, Andrea Lengerer, Julia Schroedter, Paul Lüttinger, Christof Wolf		5. Abschlussdatum des Vorhabens 31.12.2006
4b. Autoren der Publikation (Name, Vorname(n))		6. Veröffentlichungsdatum 31.05.2007
		7. Form der Publikation
8. Durchführende Institution  Zentrum für Umfragen, Methoden und Analysen (ZUMA) Postfach 12 21 55 68072 Mannheim		9. Ber. Nr. Durchführende Institution
		10. Förderkennzeichen 07SWF05B
		11a. Seitenzahl Bericht 65
		11b. Seitenzahl Publikation
13. Fördernde Institution  Bundesministerium für Bildung und Forschung (BMBF) 53170 Bonn		12. Literaturangaben
		14. Tabellen
		15. Abbildungen
16. Zusätzliche Angaben		
17. Vorgelegt bei (Titel, Ort, Datum)		
18. Kurzfassung Das Ziel des Pilotprojekts zum Aufbau eines Servicezentrums für Mikrodaten der GESIS beim ZUMA bestand in der Verbesserung des Angebotes an Serviceleistungen für die Scientific Use Files des Mikrozensus. Dies bezog sich auf drei Bereiche: Zunächst mussten für die aktuellen Erhebungen zwischen 1999 und 2004 die Aufbereitungs- und Dokumentationsarbeiten komplettiert und abgeschlossen werden. Darüber hinaus bestand eine zentrale Aufgabe in der Erweiterung des Datenangebotes an Mikrozensus-Scientific Use Files für den Erhebungszeitraum vor 1989, für den bis zum Projektbeginn noch keine Mikrozensusdaten (SUFs) für die Forschung zur Verfügung standen. Aufgrund der in den letzten Jahren stark gestiegenen Anzahl an verfügbaren Scientific Use Files und damit verbundener Arbeitsmaterialien bestand ein drittes Ziel im Aufbau eines datenbankgestützten Metainformationssystems, um die zunehmende Komplexität des Datenmaterials für den Nutzer effektiver handhabbar zu machen.		
19. Schlagwörter: Mikrozensus, Rückerschließung, Mikrodaten-Informationssystem, Scientific Use File		
20. Verlag	21. Preis	